Microbial Diversity and Function in Radionuclide Impacted Soils and Sediments

Peter John Leary

Thesis submitted for the degree of Doctor of Philosophy

School of Natural and Environmental Sciences

June 2019

Abstract

The activity of indigenous microbial communities in a natural environment can mediate the biogeochemical cycling of key nutrients and contaminants and impose conditions (E_h and pH) that are the principal controls on metal, and radionuclide, behaviour. Contemporary molecular techniques, in particular high throughput DNA sequencing techniques and subsequent bioinformatic analysis, can be used to construct high resolution microbial taxonomic and functional profiles associated with natural environments to augment and extend our understanding of these systems, when combined with measurements of inorganic and organic geochemical profiles. Understanding the assemblages and diversity of microbial communities in natural environments can enable detailed inferences of the complexity and coupled interactions of prevailing geochemical conditions and processes of an environment to be made. This approach is particularly useful for predicting the long-term transport and fate of radionuclide contaminants, *via* both direct and indirect interaction with indigenous microbial communities. This project, as part of the LoRISE consortium (Long-lived Radionuclides in the Near Surface Environment) studied depth cores from a range of radionuclide-impacted sites from around the U.K. The naturally uranium (U) contaminated site (Needle's Eye, Scotland) represented a site that exhibited a close resemblance of the classical thermodynamically-controlled vertical succession of microbial redox processes (aerobic through to anaerobic processes). Reconstructed microbial community profiles were able to anticipate the prevailing U species as a function of depth. Sellafield-impacted sites on the north west coast of England and within the Irish Sea represented more heterogeneous microbial profiles but that were still able to explain geochemical conditions observed at the sites. Further away from Sellafield, depth cores from Loch Etive (Scotland) that were anticipated to exhibit a classic redox succession profile, did not, but contained depth-related community succession focused on sulfur cycling and fermentation. The development of bioinformatics pipelines centred on the latest platforms, databases, and statistical analyses as part of this Ph.D., enabled a high level of interrogation of the DNA sequencing data that enabled specific assertions of the geochemical conditions of each site to be made based on the sequencing results, which proved to be reliable, consistent, and accurate across all sites.

Acknowledgements

Firstly, thank you to everybody in the School Formally-Known-As-CEGs, both academic and faculty, who have had such a positive and helpful impact on my degree. Yvonne Hall, Melissa Ware, and Joanne Pritchard are among the brilliant faculty staff that put up with me organising last-minute trips and answered all my silly questions and issues in the most pleasant and polite manner possible. Many thanks to the people who make up and once made up the Geomicrobiology group. I would like to especially thank: Dr. Angie Sherry, who, frankly, deserves to be a co-author on all of this for all the help and advice she has provided over the years; Dr. Julia Rosa de Rezende and Dr. China Hanson for all their help and guidance with pipeline and statistical analysis; Dr. Tania Korin for the great discussions and data; Dr. Emma Bell, Dr. Beate Christgen, and Dr. Sven Lahme for providing me with DNA positives; Dr. Obioma Mejeha for teaching me qPCR; Dr. Luiza Lessa Andrade, for, well, being you. Thank you, Prof. Ian Head, whose vast knowledge of everything in the universe was as inspiring as it was terrifying as it was incredibly helpful. Thanks to Jonny Law and Naomi Hannaford for the amazing R course. And, importantly, to Dr. Michael Sweet, who believed in me and put me on this path.

I would like to extend my infinite gratitude and appreciation to my supervisors Dr. Clare McCann and Dr. Neil Gray. Their patience, wisdom, humour, and grace are among the greatest reasons this thesis exists at all. I am not sure either of them would ever wish to repeat the experience of trying to take someone with absolutely zero chemistry knowledge and try and make a bone-fide biogeochemist out of them ever again. But I think we just about got there.

Thank you to everybody from LoRISE. To Prof. Francis Livens and Prof. Gareth Law for making LoRISE what it was academically, and to Dr. Daisy Ray, Dr. Helena Davies, Dr. Kieran Tierney, Dr. Dimitris Kosmidis, and, I suppose, Dr. Adam Fuller, for making it so fun. Additional thanks go to Daisy and Adam for all the geochemical analysis data they shared with me that is in this thesis. Thanks to Dr. Greg Cowie and the team at SAMS for the Loch Etive cores, and thanks to Dr. Margaret Graham for the access to Needle's Eye.

Thank you to my family. To my sister, Dr. Laura Kate Wolfer, for leading the way as you have our whole lives. To my Mum, Felicity, I know you're so proud. To my Dad, Kevin, and my brother, Kit. To my dearest Auntie Justine, Uncle Pat, and the whole Hopewell family, especially my main man Aidan (I miss you!). My best friends, Jake and Danielle, for always keeping me grounded! To our beloved cat, Toony, whose untimely demands for attention/food were secretly a welcome distraction.

And, most of all, to Charlotte Olivia. I am so immensely grateful for your unwavering love and support. Your unabated joy and positivity have motivated me and kept me going throughout. Thank you, truly.

Funding for this work was provided by NERC, the Environment Agency, and Radioactive Waste Management.

Contents

Abstract iii

Acknowledgements iv

Contents vi

List of Figures xii

List of Tables xxiii

Abbreviations xxv

Chapter 1. Introduction 1

1.1 Microbial Biogeochemical Cycling in Natural Environments 1

1.2 The Interaction Between Microorganisms and Radionuclides 7

1.2.1 Uranium Biogeochemistry 8

1.2.2 Plutonium Biogeochemistry 11

1.2.3 Biogeochemistry of Additional Radionuclides of Interest 12

1.3 The Natural Laboratories 13

1.3.1 The Needle's Eye Natural Analogue Site 14

1.3.2 Sellafield-Derived Radionuclides: The Esk Estuary Saltmarshes and the North East Irish Sea Basin 16

1.3.3 The Loch Etive Basins: The 'Control' Site 19

1.4 The Advent of High Throughput DNA Sequencing 21

1.4.1 Marker Gene DNA Sequencing Methods 22

1.4.2 Marker Gene Sequencing Bioinformatics Pipeline Analysis 23

1.4.3 Shotgun Metagenome Sequencing Methods 27

1.4.4 Other High-Throughput Techniques 28

1.4.5 Geochemical Analysis Methods 30

1.5 Research Aim and Objectives 32

1.6 References 32

Chapter 2. Methods and Materials 45

2.1 Sample Collection 45

2.1.1 Needle's Eye 45

2.1.2 Ravenglass and Irish Sea Mud Patch 46
2.1.3 Loch Etive 46
2.2 DNA Extraction and Quantification 47
2.3 'End Point' PCR of Functional, 16S, and ITS Genes 48
2.4 DGGE Preparation and Analysis 49
2.5 DNA Sequencing using next generation sequencing platforms 53
2.5.1 16S rRNA Gene Sequencing 53
2.5.2 Sequencing of the Fungal ITS Region 57
2.6 16S rRNA Gene Sequencing Data – Pipeline Analysis 59
2.6.1 Development of QIIME1 Pipeline and Exploration of OTU Picking Methods 59
2.6.2 ITS2 Sequencing Analysis 61
2.6.3 Moving to QIIME2 62
2.6.4 DADA2 and Amplicon Sequence Variants 64
2.6.5 Closed-Reference OTU Picking 64
2.6.6 Taxonomy Classification 65
2.6.7 Sequence Alignment and Phylogenetic Trees 65
2.6.8 Alpha and Beta Diversity Analyses 65
2.6.9 Tax4Fun 66
2.6.10 QIIME2 Pipeline Development 68
2.7 Shotgun Metagenome Sequencing Analysis 69
2.7.1 GraftM 69
2.8 Statistical Analysis and Graphics in R 70
2.8.1 Beta Diversity Ordination Plots in R 72
2.8.2 Environmental Fitting 73
2.9 Geochemical Analysis 73
2.9.1 Needle's Eye Contaminated Core X-Ray Absorption Spectroscopy 74
2.9.2 Ravenglass and Mud Patch Organic Matter, Porewater and Radionuclide Geochemistry 75
2.10 References 76

Chapter 3. Microbial Community Analysis by Next Generation Sequencing for Biogeochemical Functional Profile Inferences in a Uranium Contaminated Soil .81

3.1 Introduction 81

3.2 Methods 83

3.2.1 Sample Collection 83

3.2.2 DNA Extraction and DNA Quantification 83

3.2.3 'End Point' PCR of Functional Genes 84

3.2.4 DNA Sequencing 84

3.2.5 16S rRNA Gene Sequencing Analysis 86

3.2.6 Tax4Fun Analysis 86

3.2.7 Shotgun Metagenome Sequencing 86

3.2.8 GraftM Analysis 87

3.2.9 Geochemical Analysis 87

3.2.10 Environmental Fitting 87

3.3 Results 88

3.3.1 DNA Extractable-Quantities and Total Organic Carbon as a Function of Depth 88

3.3.2 Qualitative Assessment of Functional Gene Distributions as a Function of Depth by 'End Point' PCR 88

3.3.3 16S rRNA Gene Sequencing Analysis Results 94

3.3.4 Taxonomic Analysis 101

3.3.5 Tax4Fun Analysis Results 111

3.3.6 Geochemical Analysis Results 123

3.3.7 Environmental Fitting 128

3.3.8 GraftM Analysis Results 129

3.4 Discussion 134

3.4.1 The Contaminated Depth Core 135

3.4.2 The Control Depth Core 145

3.4.3 The Fate of Radionuclides at Needle's Eye 149

3.5 References 152

Chapter 4. Using Microbial Community and Function to Understand and Predict Biogeochemical Cycling and Radionuclide Fate in Two Sellafield-Impacted Sediment Cores from the Eastern Irish Sea 159

4.1 Introduction 159

4.2 Methods 161

4.2.1 Sample collection and preparation 161

4.2.2 DNA Extraction and Quantification 161

4.2.3 'End Point' PCR Functional Gene Series 161

4.2.4 16S rRNA Gene Sequencing Preparation 162

4.2.5 16S rRNA Gene Sequencing Pipeline Analysis 162

4.2.6 Tax4Fun Analysis 162

4.2.7 Geochemical Analysis 162

4.3 Results 162

4.3.1 DNA Extractions 162

4.3.2 Total Organic Carbon 165

4.3.3 'End Point' PCR Functional Gene Series 166

4.3.4 16S rRNA Gene Sequencing Results 169

4.3.5 Taxonomy Analysis 175

4.3.6 Tax4Fun Inferred Functional Gene Abundance Profiles 186

4.3.7 Geochemistry 199

4.3.8 Environmental Fitting of Geochemical Data 205

4.4 Discussion 207

4.4.1 Contrasting the Ravenglass and Mud Patch Depth Cores 207

4.4.2 The Ravenglass Depth Core 209

4.4.3 The Mud Patch Depth Core 215

4.4.4 The Fate of Radionuclides at Ravenglass and Irish Sea Mud Patch 219

4.5 References 221

Chapter 5. Using Next-Generation Sequencing Techniques to Investigate Microbial Diversity and Function in Sediment Cores from Loch Etive, Scotland 227

5.1 Introduction 227

5.2 Methods 229
5.2.1 Sample Collection 229
5.2.2 DNA Extraction and Quantification 229
5.2.3 'End Point' PCR Functional Gene Series 229
5.2.4 16S rRNA Gene Sequencing Preparation 230
5.2.5 16S rRNA Gene Sequencing Pipeline Analysis 230
5.2.6 Tax4Fun Analysis 230
5.2.7 Geochemical Analysis 230
5.2.8 Environmental Fitting 230
5.3 Results 230
5.3.1 DNA Extractions 230
5.3.2 'End Point' PCR Functional Gene Series 232
5.3.3 16S rRNA Gene Sequencing Results 234
5.3.4 Differential Abundance Analysis 241
5.3.5 Taxonomic Analysis 243
5.3.6 Tax4Fun Results 255
5.3.7 Geochemical Results 267
5.3.8 Environmental Fitting 274
5.4 Discussion 276
5.4.1 Comparison of General Characteristics of Microbial Community Structures at the Shallow and Deep Basin 276
5.4.2 The Shallow Basin Depth Core: Specific Taxonomic, Functional and Geochemical Features 280
5.4.3 The Deep Basin Depth Core: Specific Taxonomic, Functional and Geochemical Features 286
5.4.4 The Fate of Radionuclides in the Loch Etive Cores 289
5.5 References 291
Chapter 6. Conclusions and Discussion for Future Work 298
6.1 General Conclusions Across All Depth Cores 298
6.2 The Benefits and Limitations of Using Microbial Functional Inference Tools 301

6.3 Using Taxonomic and Inferred Functional Profiles to Predict Radionuclide Redox Fate 306
6.4 Future Work and Considerations 309
6.4 References 310
Appendix A. Needle's Eye: AOV and Tukey HSD Tables 313
Appendix B. Needle's Eye: Geochemical Plots 321
Appendix C. Needle's Eye: GraftM Heatmaps 323
Appendix D: Ravenglass and Irish Sea Mud Patch: AOV and Tukey HSD Tables 325
Appendix E. Loch Etive: AOV and Tukey HSD Tables 335

List of Figures

Figure 1.1. Three pathways of methanogenesis: hydrogenotrophic pathway (marker 2), methylotrophic pathway (3), acetoclastic pathway (4). Figure from T. Korin (2018)..... 6

Figure 1.2. Pourbaix diagram denoting the possible equilibrium phases of a set of ions. Uranium ions are denoted by the black lines, and coloured lines denote a range of terminal electron acceptors. Diagram produced in The Geochemist's Workbench by N. Gray (pers. comm.) 8

Figure 1.3. Map of the United Kingdom and the location of the four natural laboratories. From LoRISE Proposal..... 13

Figure 1.4. From Jamet *et al.* (1993). Uranium mineralisation at Needle's Eye, demonstrating the accumulation of U in humic sediments at the base of the cliff 15

Figure 1.5. From MacKenzie *et al.* (1991). BGS schematic diagram showing the main hydrological features of the Needle's Eye natural analogue site. Pitchblende vein is embedded in the Southwick Cliff 16

Figure 1.6. Map of the UK with inset of the Irish Sea and Cumbrian coastline, and red dots indicating the location of the: Irish sea mud patch; the Sellafield nuclear site, and; the Ravenglass saltmarshes. From Ray *et al.* (in prep.) 18

Figure 1.7. Map of Loch Etive with inset of map of Scotland. Contours denote 20 cm depth. Connel is the opening to sea. Bonawe is the sill between the shallow and deep basin. From Wood *et al.* (1973) 20

Figure 3.1. Contaminated Core: Depth profiles showing: (A). The amount of DNA extracted from each depth sample from the core in ng/μL, and; (B). The total organic carbon (TOC) measured by Loss on Ignition (LOI) for the depth samples measured. Point colours and shapes denote sample depth group: Blue circles = Group 1 (0 – 10 cm); Red triangles = Group 2 (11 – 20 cm); Green diamonds = Group 3 (21 – 29 cm), and; Purple squares = Group 4 (31 – 41 cm)..... 89

Figure 3.2. Contaminated Depth Core: XY scatterplot showing the log-transformed amount of DNA extracted from each sample against the log-transformed TOC value for the same sample. Spearman correlation is presented as ρ and p values..... 89

Figure 3.3. Control Core: Depth profiles showing the amount of DNA extracted from each depth sample from the core in ng/μL. Point colours and shapes denote sample depth group: Blue circles = Group 1 (0 – 10 cm); Red triangles = Group 2 (11 – 20 cm); Green diamonds = Group 3 (21 – 29 cm), and; Purple squares = Group 4 (31 – 41 cm). 90

Figure 3.4. Contaminated Core: Depth profiles of 'end point' PCR target gene band intensities, as determined by agarose gel electrophoresis and BioNumerics, as a function of depth in cm. Colour shapes refer to sample depth group. (A). Bacterial *amoA*; (B). Archaeal *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA*..... 92

Figure 3.5. Control Core: Depth profiles of 'end point' PCR target gene band intensities, as determined by agarose gel electrophoresis and BioNumerics, as a function of depth in cm. Colour shapes refer to sample depth group. (A). Bacterial *amoA*; (B). Archaeal *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA* 95

Figure 3.6. Contaminated Core: Whisker boxplots denoting range in alpha diversity indices values of the four depth sample groups from the Needle's Eye contaminated core. Colours denote depth group sample: Blue = Group 1 (0 – 10 cm); Red = Group 2 (11 – 20 cm); Green = Group 3 (21 – 29 cm), and; Purple = Group 4 (31 – 41 cm). (A). Shannon Species Richness; (B). Faith Phylogenetic Diversity; (C). Pielou Species Evenness 96

Figure 3.7. Control Core: Whisker boxplots denoting range in alpha diversity indices values of the four depth sample groups from the Needle's Eye contaminated core. Colours denote depth group sample: Blue = Group 1 (0 – 10 cm); Red = Group 2 (11 – 20 cm); Green = Group 3 (21 – 29 cm), and; Purple = Group 4 (31 – 41 cm). (A). Shannon Species Richness; (B). Faith Phylogenetic Diversity; (C). Pielou Species Evenness 97

Figure 3.8. Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity metric of samples from both depth cores. Blue denotes contaminated core samples and red denotes control core samples. Point shapes denote sampling depth groups. Black circles denote samples without a label (due to space constraints). Red outline circles denote hierarchical clustering of samples based on centroid similarity 98

Figure 3.9. Detrended correspondence analysis (DCA) of samples from both depth cores. Blue denotes contaminated core samples and red denotes control core samples. Red outline circles denote hierarchical clustering of samples based on centroid similarity 99

Figure 3.10. Contaminated Core: nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Colours and points denote sample depth groups 99

Figure 3.11. Contaminated Core: DCA plot of microbial community assemblages of samples. Colours and points denote sample depth groups 100

Figure 3.12. Control Core: nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Colours and points denote sample depth groups..... 100

Figure 3.13. Control Core: DCA plot of microbial community assemblages of samples. Colours and points denote sample depth groups 101

Figure 3.14. Contaminated Core: 100% stacked bar chart of the ten most abundant ASV taxonomies collapsed at genus level. Size of bar denotes percentage abundance of the genera as a proportion of total ASV genera assemblage within each sample. Dark blue = uncultured *Acidobacteria* subgroup 6; dark green = uncultured *Anaerolineaceae*; red = uncultured methanogenic *Bathyarchaeia*; yellow = uncultured *Chloroflexi* KD4-96; purple = *Sulfurifustis* spp.; blue = uncultured δ-*proteobacteria* Sva0485; light green = uncultured *Rokubacteriales* NC10; light blue = uncultured δ-*proteobacteria* MBNT15; gold = uncultured *Thermoprotei* archaeon,

Bathyarchaeia; orange = uncultured *Thermodesulfovibriona*, *Nitrospirae*. 103

Figure 3.15. Contaminated Core: Depth profiles of six additional key functional microbial taxa. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). Uncultured *Nitrosopumilaceae*; (B). *Candidatus Methylomirabilis* spp.; (C). uncultured *Nitrosomonadaceae* MND1; (D). *Methanosaeta* spp.; (E). *Methanolinea* spp.; (F). *Methanoregula* spp. ... 104

Figure 3.16. Control Core: 100% stacked bar chart of the ten most abundant ASV taxonomies collapsed at genus level. Size of bar denotes percentage abundance of the genera as a proportion of total ASV genera assemblage within each sample. Dark blue = uncultured *Acidobacteria* Subgroup 6; dark green = uncultured methanogenic *Bathyarchaeia*; red = uncultured *Gallionellaceae*; yellow = uncultured α-*proteobacteria*; purple = uncultured *Anaerolineaceae*; blue = uncultured *Nitrospirae* 4-29-1; green = uncultured *Acidobacteriales*; light blue = uncultured δ-*proteobacteria* MBNT15; gold = uncultured *Rokubacteria* NC10; orange = uncultured *Thermodesulfovibriona*. 109

Figure 3.17. Control Core: Depth profiles of six additional key functional microbial taxa. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). *Desulfobacca* spp.; (B). *Pseudolabrys* spp.; (C). *Sideroxydans* spp.; (D). uncultured *Micropepsaceae*; (E). *Nitrospira* spp.; (F). *Candidatus Nitrosotalea* spp. 110

Figure 3.18. Contaminated Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirS*; (C). *nirK*; (D). *dsrA*; (E). *mcrA* 112

Figure 3.19. Contaminated Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA* 113

Figure 3.20. Contaminated Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD* 115

Figure 3.21. Contaminated Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS* 118

Figure 3.22. Control Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirS*; (C). *nirK*; (D). *dsrA*; (E). *mcrA* 119

Figure 3.23. Control Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA* 120

Figure 3.24. Control Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD* 122

Figure 3.25. Control Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). NiFe hydrogenase; (F). *mtdB*; (G). *hdrA*; (H). *cooS* 124

Figure 3.26. Contaminated Core: From Fuller *et al.* (in prep.). Depth profiles of: (A). Si = ○, organic matter (OM) = ■; (B). Mn/Al ratio; (C). Fe/Al ratio; (D). U concentration; and (E). HCl extractable solid phase Fe^{2+} as a % of total extractable Fe. Porewater: (F). Fe, and (G). U concentrations. 126

Figure 3.27. Contaminated Core: Depth profile of percentage of total U determined to be U^{6+} versus U^{4+} by LCF fitting of the U L_3-edge XANES (red line, circle points) and ITFA analysis of the U M_4-edge HERFD-XANES (blue dotted line, triangles). 127

Figure 3.28. Contaminated Core: nMDS based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Environmental Fitting correlation is represented by the blue arrows. Colours and shapes denote sample depth group 128

Figure 3.29. Contaminated Core: DCA of the microbial community assemblages between samples. Environmental fitting correlation is represented by the blue arrows 129

Figure 3.30. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. From Ettwig *et al.* (2010) 145

Figure 4.1. Ravenglass Depth Core: (A). The amount of DNA extracted from each depth sample from the core in ng/μL, and; (B). The total organic carbon (TOC) measured by Loss on Ignition (LOI) for the depth samples measured. Point colours and shapes denote sample depth group: blue circles = group 1 (1 – 5 cm); red triangles = group 2 (6 – 10 cm); green diamonds = group 3 (12 – 20 cm), and; purple squares = group 4 (22 – 28 cm) 163

Figure 4.2. Ravenglass Depth Core: XY scatterplot showing the amount of DNA extracted from each sample against TOC value for the same sample. Spearman correlation is presented as ρ and p values 164

Figure 4.3. Mud Patch Depth Core: (A). The amount of DNA extracted from each depth sample from the core in ng/μL, and; (B). The total organic carbon (TOC) measured by Loss on Ignition (LOI) for the depth samples measured. Point colours and shapes denote sample depth group: blue circles = group 1 (1 – 5 cm); red triangles = group 2 (6 – 10 cm); green diamonds = group 3

(12 – 20 cm); purple squares = group 4 (22 – 30 cm); gold stars = group 5 (32 – 36 cm)..... 164

Figure 4.4. Mud Patch Depth Core: XY scatterplot showing the amount of DNA extracted from each sample against TOC value for the same sample. Spearman correlation is presented as ρ and p values..... 165

Figure 4.5. Functional gene end point PCR band intensities for the Ravenglass core. Values on the Y axis represent depth in cm, values on the X axis represent the band intensity of the PCR amplicon as calculated by BioNumerics. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA*..... 167

Figure 4.6. Functional gene end point PCR band intensities for the Mud Patch core. Values on the Y axis represent depth in cm, values on the X axis represent the band intensity of the PCR amplicon as calculated by BioNumerics. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA*..... 168

Figure 4.7. Whisker box plots of three alpha diversity indices from the Ravenglass and Mud Patch depth cores. Boxes denote samples divided into arbitrary groups in depth clusters for comparison. Group A = 0 – 5 cm; Group B = 6 – 10 cm; Group C = 12 – 20 cm; Group D = 22 – 28 cm for Ravenglass, and 22 – 30 cm for Mud Patch; Group E = 32 – 36 cm, Mud Patch only. (A). Boxplot of Shannon species richness of Ravenglass; (B). Shannon richness for Mud Patch; (C). Faith phylogenetic diversity for Ravenglass; (D). Faith phylogenetic diversity for Mud Patch; (E). Pielou species evenness for Ravenglass; (F). Pielou species evenness for Mud Patch..... 170

Figure 4.8. Ordination of the ASV taxonomic classification assemblages collapsed at species classification level of both depth core samples. Red samples denote Mud Patch depth core and blue samples denote Ravenglass depth core. Points refer to sample group depth. (A). Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity of taxonomy collapsed at species level; (B). Detrended Correspondence Analysis (DCA) of taxonomy collapsed at species level..... 173

Figure 4.9. Ravenglass Depth Core: PCoA of the microbial community assemblages. Sample colours denote grouping of depth samples. Red = Group 1 (0 – 5 cm); Green = Group 2 (6 – 10 cm); Blue = Group 3 (12 – 20 cm); Purple = Group 4 (22 – 28 cm). Labels represent depth in cm. (A). Bray Curtis similarity; (B). Weighted UniFrac similarity; (C). Unweighted UniFrac similarity..... 174

Figure 4.10. Ravenglass Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups 174

Figure 4.11. Mud Patch Depth Core: PCoA of the microbial community assemblages. Sample colours denote grouping of depth samples. Red = Group 1 (0 – 5 cm); Gold = Group 2 (6 – 10 cm); Green = Group 3 (12 – 20 cm); Blue = Group 4 (22 – 30 cm); Purple = Group 5 (32 – 36 cm). Labels

represent depth in cm. (A). Bray Curtis similarity; (B). Weighted UniFrac similarity; (C). Unweighted UniFrac similarity 175

Figure 4.12. Mud Patch Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups 175

Figure 4.13. Ravenglass Depth Core: 100% stacked bar chart of ASV taxonomy collapsed at genus level, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = *Woeseia* spp.; dark green = uncultured *Ardenticatenales* (1); red = uncultured δ-*proteobacteria* MBNT-15; yellow = uncultured *Anaerolineaceae*; purple = uncultured *Gemmatimonadetes* PAUC43f marine benthic group; blue = uncultured *Pirellulaceae*; green = uncultured γ-*proteobacteria*; light blue = *Pirellulaceae* Pir4 lineage; gold = *Thermoanaerobaculaceae* subgroup 10; orange = uncultured *Ardenticatenales* (2) 179

Figure 4.14. Ravenglass Depth Core: Depth profiles of ten additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). *Desulfobulbaceae* (1); (B). *Chloroflexi* KD4-96; (C). *Sandaracinaceae*; (D). *Dadabacteriales*; (E). *Desulfobulbaceae* (2); (F). *Saprospiraceae*; (G). *Prolixibacteraceae*; (H). *Methyloceanibacter* spp.; (I). *Nitrosopumilaceae*; (J). *Bathyarchaeia* 180

Figure 4.15. Mud Patch Depth Core: 100% stacked bar chart of ASV taxonomy collapsed at genus level, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = uncultured *Planococcaceae*; dark green = uncultured *Desulfobulbaceae* (1); red = uncultured *Flavobacteriaceae*; yellow = uncultured γ-*proteobacteria* B2M28; purple = *Psychrobacter* spp.; blue = uncultured γ-*proteobacteria* (1); light green = uncultured *Syntrophobacteraceae*; light blue = uncultured *Anaerolineaceae*; gold = uncultured γ-*proteobacteria* (2); orange = uncultured *Desulfobulbaceae* (2) 184

Figure 4.16. Mud Patch Depth Core: Depth profiles of ten additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). *Lutimonas* spp.; (B). *Woeseia* spp.; (C). *Thermoanaerobaculaceae* subgroup 10; (D). *Methyloceanibacter* spp.; (E). δ-*proteobacteria* SEEP-SRB1; (F). *Bathyarchaeia* 185

Figure 4.17. Ravenglass Depth Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA* 187

Figure 4.18. Ravenglass Depth Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is

represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA* 188

Figure 4.19. Ravenglass Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD* 189

Figure 4.20. Ravenglass Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS* 191

Figure 4.21. Mud Patch Depth Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA* 193

Figure 4.22. Mud Patch Depth Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA* 194

Figure 4.23. Mud Patch Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD* 196

Figure 4.24. Mud Patch Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS* 198

Figure 4.25. Ravenglass Depth Core: Depth profiles of porewater Mn (A)., Fe (B). concentrations in μM, and porewater sulfate (SO_4^{2-}) concentration (C). in mM from ion chromatography (IC).. 200

Figure 4.26. Ravenglass Depth Core: Depth profile of XRF Mn (A). and Fe (B). values normalised by aluminium (Al). 200

Figure 4.27. Ravenglass Depth Core: Depth profiles of activity of four radionuclide isotopes. X axis denotes activity in Bq/Kg and Y axis depth in cm. (A). ^{241}Am; (B). ^{137}Cs; (C). $^{239/240}$Pu; (D). ^{238}Pu 201

Figure 4.28. Mud Patch Depth Core: Depth profiles of porewater manganese (Mn) (A)., iron (Fe) (B). and, porewater SO_4^{2-} concentration in μM (C). as a function of depth from ion chromatography (IC) 203

Figure 4.29. Mud Patch Depth Core: Depth profiles of XRF Mn (A). and Fe (B). values normalised by aluminium (Al) 203

Figure 4.30. Mud Patch Depth Core: Depth profiles of activity of four radionuclide isotopes. X axis denotes activity in Bq/Kg and Y axis depth in cm. (A). ^{241}Am; (B). ^{137}Cs; (C). $^{239/240}$Pu; (D). ^{238}Pu 204

Figure 4.31. Ravenglass Depth Core: DCA plot with environmental variable data represented via environmental fitting correlation, represented by the blue

arrows. Direction of arrows shows direction of increasing gradient and length of arrow is proportional to the correlation between the variable and the ordination axis 205

Figure 4.32. Mud Patch Depth Core: DCA plot with environmental fitting added 206

Figure 5.1. Depth profiles showing the amount of DNA in ng/μL extracted from each sample from the Shallow Basin depth core (A). and the Deep Basin depth core (B). Colours/points refer to sample groups as defined by sampling resolution: Blue circles = group 1 (0 – 2 cm, 0.5 cm resolution); red triangles = group 2 (3 – 10 cm, 1 cm resolution); green diamonds = group 3 (12 – 20/22 cm for Shallow and Deep Basin respectively, 2 cm resolution); purple squares = group 4 (25/27 – 42/47 cm for Shallow/Deep Basin respectively, 5 cm resolution, including the final sample at 7 cm resolution from Shallow Basin)..... 231

Figure 5.2. Shallow Basin Depth Core: Functional gene 'end point' PCR band intensities for the Ravenglass core. Values on the Y axis represent depth in cm, values on the X axis represent the band intensity of the PCR amplicon as calculated by BioNumerics. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS* 233

Figure 5.3. Deep Basin Depth Core: Functional gene end point PCR band intensities for the Ravenglass core. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*..... 234

Figure 5.4. Whisker box plots of three alpha diversity indices for the Shallow Basin and Deep Basin depth cores. Boxes denote samples that were sampled at the same resolution: Group 1 = 0.5 cm; 2 = 1 cm; 3 = 2 cm; 4 = 5 cm. (A). Boxplot of Shannon richness diversity of Shallow Basin. (B). Shannon richness diversity of Deep Basin. (C). Faith phylogenetic diversity of Shallow Basin. (D). Faith phylogenetic diversity of Deep Basin. (E). Pielou species evenness of Shallow Basin. (F). Pielou species evenness of Deep Basin..... 236

Figure 5.5. (A). Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity metric of samples from both depth cores. Blue denotes Shallow Basin depth core samples and red denotes Deep Basin depth core samples. Point shapes denote sampling depth groups. Black circles denote samples without a label (due to space constraints). Red outline circles denote hierarchical clustering of samples based on centroid dissimilarity. (B). Detrended correspondence analysis (DCA) of samples from both depth cores..... 237

Figure 5.6. Ordination of the ASV taxonomic classification assemblages collapsed at genus level of both depth core samples. Blue samples denote Shallow Basin depth core, and red samples denote Mud Patch depth core. Points refer to sample depth groups. (A). Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity of taxonomy collapsed at genus level; (B). Detrended Correspondence Analysis (DCA) of taxonomy collapsed at genus level 238

Figure 5.7. Shallow Basin Depth Core: PCoA based on three dissimilatory metrics. Colours/points refer to sampling resolution groups. Red circles = Group 1 (0 – 2 cm); Green triangles = Group 2 (3 – 10 cm); Blue squares = Group 3 (12 – 20 cm); Purple crosses = Group 4 (25 – 42 cm). Numbers represent depth in cm. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac dissimilarity; (C). Unweighted UniFrac dissimilarity 239

Figure 5.8. Shallow Basin Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups 240

Figure 5.9. Deep Basin Depth Core: PCoA based on three dissimilatory metrics. Colours/points refer to sampling resolution groups. Red circles = Group 1 (0 – 2 cm); Green triangles = Group 2 (3 – 10 cm); Blue squares = Group 3 (12 – 22 cm); Purple crosses = Group 4 (27 – 47 cm). Numbers represent depth in cm. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac dissimilarity; (C). Unweighted UniFrac dissimilarity..... 241

Figure 5.10. Deep Basin Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups 242

Figure 5.11. Heatmap of individual ASV p-values that have been log transformed and dendrogram of hierarchical clustering of differentially abundant ASVs between the two depth cores. The spectrum of the heatmap indicates that an ASV that is red has a more positive negative-log p-value and thus more proportionally abundant in the respective sample, where as a blue ASV has a more negative negative-log p-value. Dendrogram denotes Ward's hierarchical clustering of ASVs that co-occur with each other. Red bars labelled 'y' denote the balances, with light red representing the numerator and dark red the denominator 243

Figure 5.12. Boxplot of the log ratio of the abundance of the taxa in 'y0' numerator (light red bar in Figure 5.X) and 'y0' denominator (dark red bar in Figure 5.11) from the Shallow Basin (top, green) and Deep Basin (bottom, orange) 244

Figure 5.13. Proportion plot of the most differentially abundant ASV taxonomies between Shallow and Deep basin depth cores of the 'y0' balance 245

Figure 5.14. Shallow Basin: 100% stacked bar chart ASV taxonomy (collapsed to genus level) abundance, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = *Tepidibacter* spp.; dark green = *Clostridiisalibacter* spp.; red = uncultured *Anaerolineaceae*; yellow = uncultured *Sandaracinaceae*; purple = uncultured *Desulfobulbaceae*; blue = *Thermoanaerobaculaceae* Subgroup 23; light green = *Fusibacter* spp.; light blue = *Bacteroidetes* BD2-2; gold = *Alkaliphilus* spp.; orange = uncultured γ-*proteobacteria* B2M28..... 248

Figure 5.15. Shallow Basin Depth Core: Depth profiles of six additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and

points denote sample depth group. (A). *Filomicrobium* spp.; (B). uncultured δ-*proteobacteria* NB1-j; (C). uncultured *Bathyarchaeia*; (D). Uncultured *Desulfobacteraceae* Sva0081; (E). *Psychrilyobacter* spp.; (F). *Thiogranum* spp..... 249

Figure 5.16. Deep Basin: 100% stacked bar chart ASV taxonomy (collapsed to genus level) abundance, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = uncultured *Anaerolineaceae*; dark green = uncultured *Sandaracinaceae*; red = *Bacteroidetes* BD2-2; yellow = uncultured *Clostridiaceae*; purple = *Thermoanaerobaculaceae* Subgroup 23; blue = uncultured *Syntrophobacteraceae*; green = uncultured *Bathyarchaeia*; light blue = *Ignavibacteriales* PHOS-HE36; gold = *Tepidibacter* spp.; orange = uncultured γ-*proteobacteria*..... 253

Figure 5.17. Deep Basin Depth Core: Depth profiles of three additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). uncultured *Desulfobulbaceae*; (B). Uncultured δ-*proteobacteria* NB1-j.; (C). *Desulfatiglans* spp.; (D). *Ignavibacterium* spp.; (E). Uncultured *Desulfobacteraceae* Sva0081; (F). Uncultured *Desulfobacteraceae* SEEP-SRB1..... 254

Figure 5.18. Shallow Basin Depth Core: Depth profile of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA* 256

Figure 5.19. Shallow Basin Depth Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*..... 257

Figure 5.20. Shallow Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD*..... 259

Figure 5.21. Shallow Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS*..... 261

Figure 5.22. Deep Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA* 262

Figure 5.23. Deep Basin: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun

(relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS* 263

Figure 5.24. Deep Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD* 264

Figure 5.25. Deep Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS* 266

Figure 5.26. Shallow Basin Depth Core: Depth profiles of porewater metal concentrations as a function of depth. (A). Manganese (Mn); (B). Iron (Fe).; (C). Uranium (U) 268

Figure 5.27. Shallow Basin Depth Core: Depth profiles of porewater anion analysis. (A). Nitrate; (B). Sulfate; (C). Thiosulfate; (D). Phosphate 268

Figure 5.28. Shallow Basin Depth Core: Depth profiles of XRF analysis presented as a percentage of abundance normalised by aluminium (Al) abundance. (A). Mn/Al; (B). Fe/Al 270

Figure 5.29. Shallow Basin Depth Core: Depth profile of ferrozine analysis showing abundance of Fe^{2+} as a percentage of $Fe^{3+/2+}$ 270

Figure 5.30. Deep Basin Depth Core: Depth profiles of porewater metal concentrations as a function of depth. (A). Mn; (B). Fe.; (C). U 271

Figure 5.31. Deep Basin Depth Core: Depth profiles of porewater anion analysis. (A). Nitrate; (B). Sulfate; (C). Thiosulfate; (D). Phosphate 272

Figure 5.32. Deep Basin Depth Core: Depth profiles of XRF analysis presented as a percentage of abundance normalised by Al abundance. (A). Mn/Al; (B). Fe/Al 273

Figure 5.33. Deep Basin Depth Core: Depth profile of ferrozine analysis showing abundance of Fe^{2+} as a percentage of $Fe^{3+/2+}$ 274

Figure 5.34. Shallow Basin Depth Core: Detrended correspondence analysis (DCA) of Shallow Basin microbial community assemblage with environmental variables vectors overlaid (blue arrows). Length and direction of the arrows relates to correlation between geochemical variable and DCA vectors. Red outline circles denote hierarchical clustering of samples based on centroid dissimilarity. 275

Figure 5.35. Deep Basin Depth Core: Detrended correspondence analysis (DCA) of Deep Basin microbial community assemblage with environmental variables vectors overlaid (blue arrows). Length and direction of the arrows relates to correlation between geochemical variable and DCA vectors. Red outline circles denote hierarchical clustering of samples based on centroid dissimilarity 275

List of Tables

Table 1.1. A subset of the common chemoheterotrophic redox pairs (oxidised and reduced species respectively) and their corresponding reduction potential (in V) from Thauer *et al.* (1977) and Madigan *et al.* (2018). E_0' at pH 7. $\Delta G^{0\prime}$ values from Lam and Kuypers (2011) 2

Table 1.2. Energy yields from the oxidation of various inorganic electron donors. From Madigan *et al.* (2018) 3

Table 1.3. Nitrogen metabolism reference pathways from KEGG. Shows the genes, the compound the gene uses, and the oxidation state of the compound of a range of nitrogen catabolic reactions. Arrows denote direction of pathway 4

Table 1.4. Dissimilatory sulfate reduction and oxidation pathway from KEGG. This pathway is reversible. 4

Table 2.1. List of functional and marker genes targeted by 'end point' PCR, description of the gene function, name of the primers, sequences of the primers from 5` to 3`, the annealing site where available, the target fragment length, the PCR cycling conditions, and the source of the primers and also PCR cycle conditions where appropriate 50

Table 2.2. List of PCR functional gene procedural positive DNA sources 51

Table 2.3. Tables comparing number of unique OTUs and ASVs identified from the same 16S rRNA gene sequencing data using a variety of methods. Table 3A shows the numbers of OTUs that were identified, and from how many reads, by open-reference OTU picking in QIIME 1.9.1 by UCLUST, VSEARCH, and SortMeRNA OTU picking methods, using the SILVA128 reference database. OTU picking was also performed at 97% and 99% similarity thresholds for OTU clustering of reads. Table 3B shows the numbers of OTUs within the tables that remained after all *de novo* OTUs had been filtered out, leaving only closed-reference OTUs. Table 3C shows the equivalent numbers derived from ASV identification in QIIME2, which does not use a reference database and only clusters identical sequences (*i.e.*, no similarity threshold) 62

Table 2.4. A list of the functional genes of interest, and their respective KEGG Ortholog numbers, that were identified using the Tax4Fun functional profiling of 16S rRNA gene sequences. 67

Table 2.5. List of genes targeted by the GraftM program. All packages were downloaded from https://data.ace.uq.edu.au/public/graftm/7/ except genes with * which were created using the Create tool within GraftM, as described in the main text 71

Table 3.1. List of functional and marker genes targeted by 'end point' PCR, description of the gene function, name of the primers, sequences of the

primers from 5` to 3`, the annealing site where available, the target fragment length, the PCR cycling conditions, and the source of the primers and also PCR cycle conditions where appropriate 85

Table 3.2. A list of functional genes that were targeted using precompiled reference packages in GraftM 87

Table 3.3. Contaminated Core: Uranium oxidation state variation with sample depth as determined by LCF fitting of the U L_3-edge XANES and ITFA analysis of the U M_4-edge HERFD-XANES. All percentages are normalised to 100 127

Abbreviations

acsB	Acetyl CoA synthase
amoA	Ammonia monooxygenase subunit A
ANME	Anaerobic methane-oxidising archaea
AOA	Ammonia-oxidising archaea
AOB	Ammonia-oxidising bacteria
AOV	One-way analysis of variance
aprA	Adenylylsulfate reductase subunit A
ASV	Amplicon Sequence Variant
ATP	Adenosine Triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pair
cooS	Anaerobic carbon-monoxide dehydrogenase
DCA	Detrended Correspondence Analysis
DGGE	Denaturing Gradient Gel Electrophoresis
DNA	Deoxyribonucleic Acid
DNRA	Dissimilatory nitrate reduction
dsrA	Dissimilatory sulfite reductase subunit A
ED	Electron Donor
Eh	Reduction potential
EXAFS	Extended X-Ray Absorption Fine Structure
FNR	Fumarate and nitrate reductase
HERFD-XANES	High Energy-Resolution Fluorescence Detection
hydB	NiFe hydrogenase (cytochrome c3 hydrogenase)
ICP-MS	Inductive Coupled Plasma
ITS	Internal Transcribed Spacer
LOI	Loss on Ignition
mcrA	Methyl coenzyme M reductase subunit A
mtdB	Methylene-tetrahydromethanopterin dehydrogenase
mmoX	Methane monooxygenase
MS	Mass spectrometry
mxaF	Methanol dehydrogenase

narG	Nitrate reductase
napA	Nitrate reductase
NCBI	National Center for Biotechnology Information
nifD	Nitrogenase
nirK	Cytochrome cd1 nitrite reductase
nirS	Copper nitrite reductase
nMDS	Non-Metric Multidimensional Scaling
NMR	Nuclear magnetic resonance
norB	Nitric oxide reductase
nosZ	Nitrous-oxide reductase
nrfA	Nitrite reductase
nt	Nucleotide
PERMANOVA	Permutational analysis of variance
pH	Negative logarithm of the concentration of hydrogen ions
PICRUSt	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
RDP	Ribosomal Database Project
RNA	Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid
PCR	Polymerase Chain Reaction
KEGG	Kyoto Encyclopedia of Genes and Genomes
OTU	Operational Taxonomic Unit
PCoA	Principal Coordinate Analysis
QIIME	Quantitative Insights into Microbial Ecology
TEA	Terminal Electron Acceptor
TOC	Total organic carbon
XRF	X-Ray Fluorescence

Chapter 1. Introduction

1.1 Microbial Biogeochemical Cycling in Natural Environments

The growth of bacteria and archaea in the environment is predicated by their ability to produce energy in the form of ATP. Catabolic reactions, involving the oxidation of a low-potential electron donor coupled to the reduction of a high-potential electron acceptor along electron transport chains, forms the basis of all energy generation in living organisms. In biological systems, catabolic reactions are divided into oxidative phosphorylation (respiration), which requires separate electron donors and acceptors, and fermentation, which solely relies on internally-balanced oxidation-reduction reactions of organic compounds. Both are defined by the amount of energy each process produces, in the form of $\Delta G^{0`}$ (Gibb's free energy, usually expressed as kJ mol^{-1}), where the more negative a redox reaction is, the more likely it is to happen spontaneously, *i.e.*, exergonic. The thermodynamic favourability of catabolic reactions is determined by the reduction potential of the reactants, denoted as $\Delta E_0'$. Reduction potential is proportional to Gibb's free energy. The strongest oxidant, molecular oxygen (O_2), has a positive reduction potential (E_0' = + 0.82 (V)), *i.e.*, more readily accepts electrons (e^-) from electron donors, producing more Gibb's free energy per mol (Table 1.1).

Microorganisms are able to use a wide range of compounds as electron donors and electron acceptors during catabolism. Microorganisms that use organic compounds as carbon and electron sources are known as chemoheterotrophs, while microorganisms that use inorganic electron donors are chemolithotrophs, and microorganisms that use inorganic carbon are chemoautotrophs. The reduction of O_2 as a terminal electron acceptor (TEA) is aerobic respiration, and a range of other compounds can be used in lieu of O_2, called anaerobic respiration. The thermodynamic favourability, that is, the amount of energy produced by the reduction of the TEA, is a key determiner in the preference for respiration and the TEA used by microorganisms (Madigan *et al.*, 2018).

The 'classic' model of chemoheterotrophic microbial respiration in natural environments is dictated by the availability of TEAs and their thermodynamic favourability. This model is defined by the stratification of TEA and redox succession as a function of depth, defined by the vertical diffusion of TEAs from the surface downwards (Thauer *et al.*, 1977). In a closed vertical system, where O_2 diffuses from the surface downwards, microorganisms will consume O_2 during aerobic respiration.

At a sufficient depth, the O_2 will be depleted, and microorganisms will be required to use alternative TEAs to continue dissimilatory respiratory growth. The succession of TEAs is defined by the $\Delta G^{0\prime}$ of the reaction, *i.e.*, the thermodynamic favourability of the TEA. As TEAs are reduced and depleted with depth, the reduction potential of the system becomes more reducing, since reduction potential is proportional. This succession of TEA favourability is shown in Table 1.1. The succession of TEA redox pairs is, in a closed vertical diffusion-limited system, inextricably linked to depth, *i.e.*, with increasing depth in a system, the transition from aerobic O_2 reduction to anaerobic respiration is expected to be observed (Thauer *et al.*, 1977; Froelich, 1979; Berner, 1980).

Table 1.1. A subset of the common chemoheterotrophic redox pairs (oxidised and reduced species respectively) and their corresponding reduction potential (in V) from Thauer *et al.* (1977) and Madigan *et al.* (2018). $E_0{}'$ at pH 7. $\Delta G^{0\prime}$ values from Lam and Kuypers (2011).

Redox Pair	$E_0{}'$	$\Delta G^{0\prime}$ (kJ per reaction)	Process	General Conditions
$O_2 \rightarrow H_2O$	+0.82	–402	Aerobic respiration	Obligate and facultative aerobes
$Mn^{4+} \rightarrow Mn^{2+}$	+0.798	–385	Manganese respiration	Facultative aerobes
$NO_3^- \rightarrow N_2$	+0.43	–355	Denitrification	Facultative aerobes
$NO_3^- \rightarrow NO$	+0.36	–371	Nitrate respiration	Facultative aerobes (some reduce nitrate to ammonia)
$Fe^{3+} \rightarrow Fe^{2+}$	+0.77	–241	Iron respiration	Facultative aerobes and obligate anaerobes
$SO_4^{2-} \rightarrow HS^-$	–0.217	–43.8	Sulfate respiration	Obligate anaerobes
$CO_2 \rightarrow CH_4$	–0.24	–19.9	Carbonate respiration	Methanogenesis, obligate anaerobes
$CO_2 \rightarrow CH_3–COO^-$	–0.30		Carbonate respiration	Acetogenic bacteria, obligate anaerobes

In addition to chemoheterotrophic processes with organic carbon as the electron donor, microorganisms are capable of performing a range of chemolithotrophic processes that utilise inorganic carbon as the electron donor. These are usually oxidative processes involved in the generation of energy linked to the oxidation of reduced compounds in aerobic conditions. A list of commonly occurring chemolithotrophic reactions in natural environments is presented in Table 1.2.

Table 1.2. Energy yields from the oxidation of various inorganic electron donors. From Madigan *et al.* (2018).

Electron Donor	Chemolithotrophic Reaction	E_0' of couple (V)	$\Delta G^{0\prime}$ (kJ per reaction)	Number of electrons per reaction	$\Delta G^{0\prime}$ (kJ/2 e^-)
Hydrogen	$H_2 + ½\ O_2 \rightarrow H_2O$	–0.42	–237.2	2	–237.2
Sulfide	$HS^- + H^+ + ½\ O_2 \rightarrow S^0 + H_2O$	–0.27	–209.4	2	–209.4
Sulfur	$S^0 + 1½\ O_2 + H_2O \rightarrow SO_4^{2-} + 2\ H^+$	–0.20	–587.1	6	–195.7
Ammonium	$NH_4^+ + 1½\ O_2 \rightarrow NO_2^- + 2\ H^+ + H_2O$	+0.34	–274.4	6	–91.6
Nitrite	$NO_2^- + ½\ O_2 \rightarrow NO_3^-$	+0.43	–74.1	2	–74.1
Ferrous iron[a]	$Fe^{2+} + H^+ + ¼\ O_2 \rightarrow Fe^{3+} + ½\ H_2O$	+0.20	–32.9	1	–65.8

[a] E_0' values are for pH 7 at 25°C. At pH 2, the E_0' for the Fe^{2+}/Fe^{3+} couple is about +0.77 V.

Aerobic microorganisms are defined by the amount of oxygen required for survival and growth. Obligate (strict) aerobes require oxygen for catabolic processes, while facultative aerobes (or microaerophiles) are able to grow in environments with very low oxygen concentrations. Many microbial groups can switch to alternative, though still relatively electropositive, TEAs. Facultative anaerobic microorganisms switch to alternative TEAs or fermentation in the absence of oxygen for catabolic growth, characterised by a switch in gene regulation such as the fumarate and nitrate reductase (FNR) transcription factor (Spiro and Guest, 1990; Unden *et al*., 1995). Obligately anaerobes cannot grow in the presence of oxygen, of which most are generally associated with electronegative respiratory processes (Madigan *et al.*, 2018). Microorganisms are also generally defined by which specific TEAs they utilise for dissimilatory energy conservation and growth. Specific microbial phylogenetic groups perform specific chemohetero- or chemolithotrophic catabolic functions. Therefore, the presence of certain microbial phylogenetic groups with specific known catabolic function can indicate the presence of these processes in an environmental system.

Microbial respiration of organic and inorganic compounds is a key driver in the cycling of a range of essential nutrients in soils and sediments around the globe, including carbon (C) (Johnston *et al.*, 2004), nitrogen (N) (Hayatsu *et al.*, 2008), sulfur (S) (Klotz *et al.*, 2011), and heavy metals including iron (Fe), manganese (Mn), and actinides (Lloyd, 2003) (discussed below).

Table 1.3. Nitrogen metabolism reference pathways from KEGG. Shows the genes, the compound the gene uses, and the oxidation state of the compound of a range of nitrogen catabolic reactions. Arrows denote direction of pathway.

Oxidation State	+5	+3	+2	+1	0	-1	-2	-3
Dissimilatory nitrate reduction	*narGHI napAB*	*nirBD nrfAH*			→			
	NO_3^-	NO_2^-			→			NH_3
Denitrification	*narGHI napAB*	*nirK nirS*	*norBC*	*nosZ*				
	NO_3^-	NO_2^-	NO	N_2O	N_2			
Nitrogen fixation					*nifDKH anfG*	→		
					N_2	→		NH_3
Nitrification		*nxrAB*			←	*Hao*	←	*amoCAB*
	NO_3^-	NO_2^-			←	NH_2OH	←	NH_3

Table 1.4. Dissimilatory sulfate reduction and oxidation pathway from KEGG. This pathway is reversible.

Oxidation State	+6		+4	+2	-2
Dissimilatory sulfate reduction and oxidation	*Sat*	*aprAB*	*dsrAB*		
	SO_4^{2-}	APS	SO_3^{2-}		HS^-

Microorganisms transform a range of N compounds as part of catabolic reactions. Pathways including nitrification, denitrification, dissimilatory nitrate reduction, nitrogen fixation, and anaerobic ammonia oxidation are among the key microbial catabolic pathways that form part of the global N cycle (Hayatsu *et al.*, 2008). Tables 1.1 and 1.2 give examples of anaerobic and aerobic N-based catabolic reactions respectively. Due to the relatively high reduction potential and energy yielded from many nitrate-reduction pathways, it is a ubiquitous pathway in many aerobic and microaerobic soils and sediments globally (Hayatsu *et al.*, 2008; Lam and Kuypers, 2011). Examples of microbial catabolic nitrogen pathways are presented in Table 1.3. Sulfur is another key compound related to microbial biomass growth, reducing and oxidised a range of sulfur compounds as part of aerobic and anaerobic respiration (Jørgensen, 1982). A range of reduced sulfur compounds can be used by bacteria as TEAs during anaerobic respiration, including sulfate, thiosulfate, sulfite, and sulfide, while other bacteria can oxidise these reduced species as part of aerobic respiration (Fike *et al.*, 2015). The reversible dissimilatory sulfate reduction and oxidation pathway is presented in Table 1.4. Methanogenesis is the process of degrading both organic and inorganic single carbon compounds into

CH_4 and is generally limited to very anaerobic and reducing conditions and the exhaustion of all other more thermodynamically favourable TEAs (Heddrich and Whitman, 2006). Methanogenesis is a complicated process that can utilise three main products, H_2, methanol, and acetate, with examples of these pathways presented in Figure 1.1 demonstrating the wide range of enzymes and intermediate products involved in multiple methanogenic pathways. Anaerobic carbon fixation, in the form of the reductive acetyl-CoA pathway (known as the Wood-Ljungdahl pathway) is the fixation of CO_2 *via* condensation reactions with H^+ to form acetate (Ragsdale and Pierce, 2008). This is a strictly anaerobic process with a low reduction potential, and commonly, acetogenic microorganisms compete with hydrogenotrophic methanogenic archaea for H^+ (Fu *et al.*, 2018) or form syntrophic partnerships with acetoclastic and methylotrophic methanogens (Ragsdale and Pierce, 2008).

Characterisation of microbial community profiles within soils and sediments, and the abundance and transition of key functional monophyletic microbial taxonomic groups that perform specific respiratory processes, as a function of depth, implies the presence of that particular catabolic function. The presence and abundance of key functional microbial taxonomic and functional groups implies that geochemical conditions supported the growth of these groups. Using microbial DNA as a biomarker to infer geochemical conditions and biogeochemical function of a site offers an alternative approach to directly measuring conditions such as pH and E_h, which can vary drastically on short spatial or temporal scales, while DNA in natural soils and sediments has a turnover period on a scale of months, giving a longer-term picture of the environmental conditions.

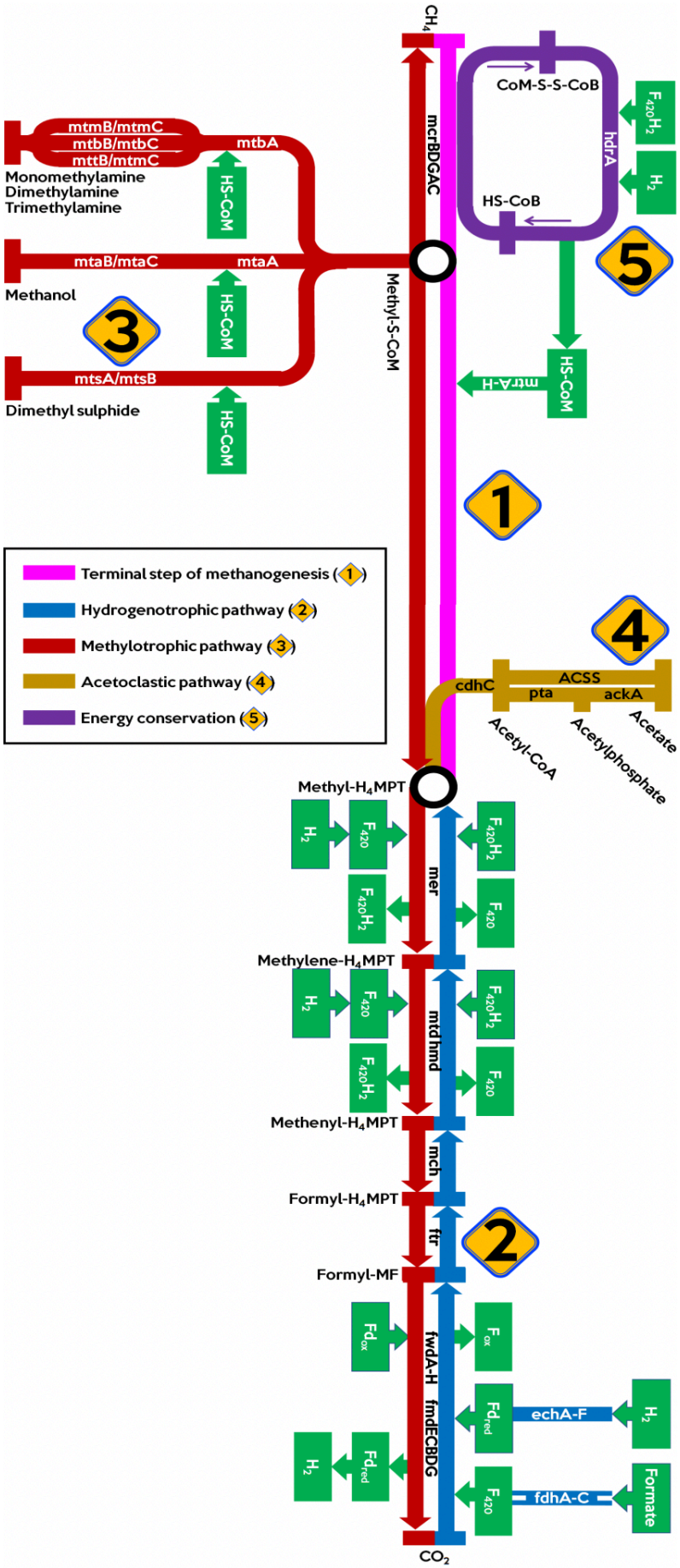


Figure 1.1. Three pathways of methanogenesis: hydrogenotrophic pathway (marker 2), methylotrophic pathway (3), acetoclastic pathway (4). Figure from T. Korin (2018)

1.2 The Interaction Between Microorganisms and Radionuclides

Radionuclides, either from geogenic sources such as U, or anthropogenic sources from nuclear activity, pose a risk to health due to their radiotoxicity and relatively long half-lives (Saenen *et al.*, 2013), and the remediation of radionuclide contaminated sites remains an area of intense focus today. The ability for microorganisms to couple the reduction of heavy metals and radionuclides to dissimilatory growth was first discovered in the late 1980s with the seminal work of Lovley *et al.* (1991, 1993). This revealed the potential to employ microorganisms in the bioremediation of radionuclide contaminated soils and sediments.

The primary radionuclides of concern are the actinides, as not only do they represent the most abundant radionuclide contaminants from nuclear weapon development and civil nuclear power generation (in the case of U and Pu) but they are also redox active (Lloyd, 2003) and are the most radiotoxic for the first 500 years of disposal (Ewing, 1999). The actinides are soluble when oxidised and insoluble when reduced, and reduction can be direct via microbially-mediated bioreduction or indirect such as pH changes (Francis, 1998). The redox species of the actinides range from +6 to +3, though not every single valency is found naturally in the environment (Lloyd *et al*., 2007). Higher valency states, principally the oxidised +6/+5 states, generally result in the formation of dioxo-cations, e.g., MO_2^{+} and MO_2^{2+}. These cations are more soluble than their reduced counterparts, due to lower charge densities and their complexation with ligands, resulting in neutral and anionic complexes (Lloyd *et al*., 2007). Reduced actinide species form simple cations, e.g., M^{3+} and M^{4+}, which with higher charge densities, are hydrolysed and lost easily from solution by reaction with surfaces, *e.g.*, UO_2 (Lloyd *et al*., 2007).

Figure 1.2 is a Pourbaix diagram created in The Geochemist's Workbench using the ACT2 redox-pH diagram drawing package with the base species set as Uranium (U++++) with an activity of 10^{-10} Mol/L and HCO_3^{-} with an activity of 10^{-3} Mol/L under standard conditions of pH 7, 25°C, pressure of 1 bar. A series of indicative phase transition lines for a range of key microbially-mediated redox couples were superimposed over the U phase diagram (Neil Gray, pers. comm.). This plot indicates the relationship between U redox speciation and microbially mediated redox couples, and shows that U would be reduced from the oxidised U^{6+} (from the uranyl ion, UO_2^{2+}, bound to a ligand) to the reduced U^{4+} (in the form of

mineralised uraninite, UO_2) at an E_h lower than Fe^{3+} reduction but higher than SO_4^{2-} reduction.

1.2.1 Uranium Biogeochemistry

U generally exists in natural in environments in either the oxidised U^{6+} or reduced U^{4+} species, with little evidence of the pentavalent U^{5+} intermediate or further-reduced U^{3+} occurring naturally (Renshaw *et al*., 2005; Brookshaw *et al*., 2012). In oxic conditions, U^{6+} is generally present as the relatively soluble uranyl ion (UO_2^{2+}) due to hydrolysis and formation of carbonate complexes at standard conditions (Langmuir, 1978; Lloyd *et al*., 2007; Belli *et al*., 2015). In reducing conditions, U^{4+} is prevalent, precipitating as the insoluble crystalline uraninite UO_2 (Lovley *et al*., 1991; Lloyd *et al*., 2007; Burgos *et al*., 2008) or non-crystalline or monomeric U^{4+} species (Bernier-Latmani *et al*., 2010) particularly in the presence of phosphate (Newsome *et al*., 2015). Reduced U^{4+} species are generally insoluble under reducing conditions but are susceptible to reoxidation in the presence of O_2 (Cerrato *et al*., 2013) and nitrate (Newsome *et al*., 2015). The process of promoting microbially-mediated bioreduction of soluble U^{6+} to insoluble U^{4+} forms the basis of *in situ* bioremediation of U in natural environments, aiming to remove porewater U which can pose a risk to health (Anderson *et al*., 2003).

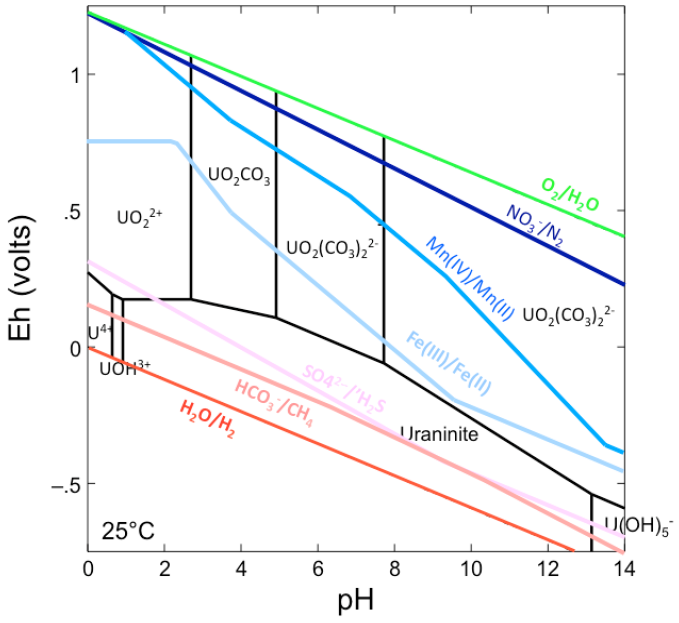


Figure 1.2. Pourbaix diagram denoting the possible equilibrium phases of a set of ions. Uranium ions are denoted by the black lines, and coloured lines denote a range of terminal electron acceptors. Diagram produced in The Geochemist's Workbench by N. Gray (pers. comm.).

Indirect reduction of U in natural environments, *e.g.*, abiotic reduction, is another consideration in bioremediation. Inorganic reductants including H_2, H_2S, and Fe^{2+}_{aq} efficiently reduce U^{6+} when it is not complexed (Suzuki and Suko, 2006). Negatively charged $UO_2(CO_3)_3^{4-}$ complexes that are common in natural environments can prevent abiotic reduction of U^{6+} to U^{4+}, although bioreduced Fe^{2+} monohydroxo surface complexes can reduce $UO_2(CO_3)_3^{4-}$ (Suzuki and Suko, 2006). Bioreduced soluble Fe^{2+} was shown to reduce U^{6+} spontaneously, especially in alkaline conditions (Du *et al*., 2011). Iron sulfides, a by-product of sulfate reduction in bacteria, were shown to abiotically reduce U^{6+} to U^{4+} albeit into the solution phase (Lee *et al*., 2013). The presence of Fe^{2+} and sulfides are thought to induce a buffering effect, preventing U^{4+} reoxidation (Newsome *et al*., 2014).

The reoxidation of biotically or abiotically reduced and insoluble U^{4+} to soluble U^{6+} is a key consideration in bioremediation efforts. Bioreduction of U^{6+} to U^{4+} in normal anaerobic conditions in the presence of carbon usually results in precipitation of U^{4+} as the crystalline uraninite (UO_2) or nanoparticulate uraninite, or the amorphous mononuclear U^{4+}, all of which are susceptible to reoxidation (Singh *et al*., 2014). Reduced U^{4+} is readily suspectable to abiotic reoxidation to soluble U^{6+} such as by the addition of O_2, Fe^{3+}, and Mn^{4+}, which are kinetically fast reactions (Suzuki and Suko, 2006), or by denitrification products such as NO_2^- (Singh *et al*., 2014). Recent work has investigated adding glycerol phosphate to stimulate bioreduction of U^{6+} to precipitate a recalcitrant crystalline U^{4+} phosphate mineral instead of uraninite (Newsome *et al*., 2015). Alternatively, the binding of U^{6+} to organic ligands can complicate bioreduction of U^{6+} to U^{4+}, for example, if the ligands are comprised of refractory or inorganic carbon, it can make the U inaccessible to microorganisms, but also insoluble and immobile (Belli *et al.*, 2015; Francis and Dodge, 2008; Gu *et al.*, 2005), potentially negating the need for a bioreduction regime.

Bacteria that can enzymatically dissimilatory reduce U were first discovered in the Fe-reducing δ-*proteobacteria* genus *Geobacter* spp. and in the γ-*proteobacteria* genus *Shewanella* spp. (Lovley *et al.*, 1991). Other bacterial species have been demonstrated to be able to reduce U, including a *Clostridium* sp., and sulfate-reducing *Desulfovibrio desulfuricans* and *D. vulgaris* (Payne *et al.*, 2002), and the *Anaeromyxobacter* spp. and *Desulfosporosinus* spp. (Cardenas *et al.*, 2010), in addition to fermentative bacteria (Francis *et al*., 1994). U^{6+} reduction is

thermodynamically similar in energy yield to Fe^{3+} reduction in standard conditions, and therefore under reducing conditions in the presence of organic matter, bacteria are anticipated to use U^{6+} as an electron acceptor during anaerobic respiration (Newsome *et al*., 2014). The exact enzymatic process of microbial U reduction is largely undescribed, but the role of *c*-type cytochromes and the transfer of e^- to *in vivo* U has been suggested (Lovley *et al.*, 1993; Payne *et al.*, 2002) as well as the role of pili and the transfer of e^- to extracellular U (Cologgi *et al.*, 2011; Lovley, 2011; Newsome *et al*., 2014), although this mechanism is still disputed (Williams *et al.*, 2013).

In addition to direct enzymatic bioreduction of U by microorganisms, alternative methods to retard the migration of oxidised U^{6+} ions are employed by microorganisms. Review of these methods are in Francis (1998) and Newsome *et al*. (2014). Biomineralisation is the process whereby U precipitates with microbially-generated ligands or carbonates on cell surfaces in alkaline conditions. Bioaccumulation is the method whereby U is actively taken up into the cell, such as in the form of uranyl phosphates, possibly in order to mitigate metal- or radiotoxicity. Biosorption is the passive uptake of U to the surface of microbial cells, but its applicability to bioremediation is limited.

Previous studies into the microbial ecology of U contaminated sites has primarily focused on sites with very highly elevated levels of U, *e.g.*, U mine tailings (Yan and Luo, 2014; Sutcliffe *et al.*, 2017; Yan *et al.*, 2016; Rastogi *et al.*, 2010). One particular study of a site with U contamination as high as 4,000 pm demonstrated that at U concentrations of < 1,500 ppm, no significant changes to microbial communities were observed, and above 1,500 ppm, significant changes to nitrogen cycling and methanogenic communities were observed (Sutcliffe *et al.*, 2017). Another study looked at soils with pH ranging from 3.7 – 6.7 and found that geochemical conditions was imposing community selection (Akob *et al.*, 2006). A more recent study that used high-throughput metagenomic sequencing of U contaminated sites of ~ 40 ppm, with a very low pH (~ pH 3) found that amino acid metabolism and signal transduction annotated genes were significantly lower in the contaminated samples and that *Actinobacteria* drove the microbial community assemblage of the contaminated samples (Yan *et al.*, 2016). It is difficult to draw accurate comparisons between these studies and the data generated in this project as other geochemical variables are likely to be very different, such as pH, geographical and hydrological features, and

the levels of contamination, which are equally going to impose on the microbial community structure in a demonstrable way.

1.2.2 Plutonium Biogeochemistry

Plutonium (Pu) is another important source of nuclear fuel in power generation produced by neutron activation of U (Newsome *et al*., 2014). Pu biogeochemistry is more complex than U, where Pu can be found in four main oxidation states at standard E_h and pH: +3, +4, +5, and +6 (Morse and Choppin, 1991), and in natural waters, can exist in multiple oxidation states simultaneously (Ewing *et al*., 2010).

The dominant oxidation species found in most natural environments is Pu^{4+}, usually in the form of the highly insoluble hydrous oxide $Pu(OH)_4$ (Newsome *et al*., 2014). Previous studies showed that only 0.01% of Pu released into soils was in solution under reducing conditions, indicating that Pu in most reducing conditions is insoluble Pu^{4+} rather than soluble oxidised 5+ or reduced 3+ (Kimber *et al*., 2012; Kimber *et al*., 2015). Pu is most stable in as soluble Pu^{5+} in aerobic conditions, the most soluble state being the plutonyl cation PuO_2^+ (Newsome *et al*., 2014). $Pu^{5+/6+}$ form dioxo-cations (*e.g.*, PuO_2^+ and PuO_2^{2+}) which remain mobile in solution. Bacteria have been shown to reduce Pu from Pu^{6+} down to Pu^{3+} *via* the 5+ and 4+ species, using it as a terminal electron acceptor (TEA) during anaerobic respiration (Boukhalfa *et al.*, 2007; Francis, 2007; Renshaw *et al.*, 2009). Pu^{4+} is expected to reduce spontaneously at around the same redox potential as Mn^{4+} reduction, and to Pu^{3+} at the same redox potential as sulfate reduction (Lloyd, 2003). Pu^{3+} is more soluble than Pu^{4+}, and is found to spontaneously reoxidise to Pu^{4+} once dissolved by an unknown oxidant (Rusin *et al.*, 1994), so where reduction of U is preferential as a retardation method, this is not necessarily the case for Pu, where further reduction of the stable, insoluble Pu^{4+} to the soluble Pu^{3+} under reducing conditions is unfavourable from a mobility perspective (Lloyd, 2003). However, one study found that during microcosm experiments that attempted to elicit anaerobic microbial Pu^{4+} reduction to Pu^{3+}, only a 'minor' increase in the mobile Pu fraction (*i.e.*, reduced Pu^{3+}) was observed in the middle of the experiment, and was removed by the end of experiment, indicating that Pu is highly refractory (Kimber *et al*., 2012).

Additionally, Pu has been shown to be associated with Fe/Mn oxide phase of lake and coastal sediments and adsorbs readily on synthetic Fe oxides (Malcolm *et*

al., 1990). Microbially-mediated redox cycling of Fe and Mn in these environments may release Pu where it will be susceptible to redox cycling (Malcolm *et al*., 1990). More recent studies of Pu in natural environments showed that between 63.8 – 94.5% of Pu was in a highly refractory residual phase in two UK soils and was mostly unsusceptible to leaching from microbially-produced sulfuric acid (Kimber *et al*., 2015).

1.2.3 Biogeochemistry of Additional Radionuclides of Interest

Neptunium (Np) is an actinide that is a by-product of Pu and americium (Am) decay from the nuclear fuel cycle. The biogeochemistry of Np is similar to that of U and Pu. In oxic conditions, Np^{5+} forms the neptunyl cation NpO_2^+ which is the most soluble of the oxidised actinides, while in reducing conditions, the poorly soluble Np^{4+} persists, primarily as NpO_2 (Lloyd *et al*., 2007). Bacteria such as *Shewanella* are able to anaerobically bioreduce Np^{5+} to Np^{4+} (Newsome *et al*., 2014).

Americium (Am) is an actinide produced by neutron activation of Pu. Reduced Am^{3+} is the most prevalent species in environmental conditions, and does not undergo redox transformations, has limited solubility and mobility, and readily sorbs to soils and sediments (Newsome *et al*., 2014). It is therefore unlikely to have any interaction with indigenous microbial populations from a redox perspective (albeit potentially in a toxicity perspective at sufficient concentrations).

Technetium is the lightest element to have isotopes that are all radioactive and is produced during nuclear fission. In aerobic conditions, the oxidised, highly soluble Tc^{7+} pertechnetate ion (TcO_4^-) persists, where it is also bioavailable as an analogue for sulfate (McBeth *et al*., 2007). Similar to the actinides, in reducing conditions, Tc^{4+} precipitates as strongly sorbing hydrous TcO_2 (Lloyd *et al*., 2007; Newsome *et al*., 2014). Also similar to the actinides, Tc can be indirectly reduced by Fe^{2+}, and due to low concentrations of Tc^{7+} in natural environments, this is the most likely mechanism of Tc reduction (McBeth *et al*., 2007). Bioreduction of Tc is also similar to that of the actinides in that it is linked to anaerobic respiration, by known iron and sulfate reducing bacteria (Newsome *et al*., 2014).

Cesium (Cs) and strontium (Sr) are alkaline metals that normally have stable isotopes in the natural environment, but radioactive isotopes of these elements are released as nuclear fission products. Neither of these elements are redox active, and

therefore not likely to interact with indigenous microbial populations. Both Cs and Sr are highly soluble and adsorb readily to clays such as bentonite (Bostick *et al*., 2002; Galamboš *et al*., 2010). Cs and Sr readily sorb to soil particles (Chiang *et al*., 2010) and can change the pH, and produce soluble ligands or new biomineral phases, in soils (Newsome *et al*., 2014).

1.3 The Natural Laboratories

Five natural environmental sites, with varying degrees of both anthropogenic and geogenic sources of radionuclide contamination and encompassing a range of habitat types including terrestrial, estuarine, and marine, were studied as each site represented a 'natural laboratory', *i.e.*, natural environments that have been altered in a detectable way from surrounding natural environments. These sites were chosen

Figure 1.3. Map of the United Kingdom and the location of the four natural laboratories. From LoRISE Proposal.

due to their elevated levels of radionuclides of interest from both naturally occurring uranium mineralisation, or authorised discharges from nuclear sites. These radionuclides contaminants have been present in the natural laboratories for at least several decades, and their distribution and cycling within the natural environments can have implications on indigenous microbial populations as well as to human health. Four of these sites were chosen as part of the LoRISE consortium (Long-lived Radionuclides in the Surface Environment – Mechanistic Studies of Speciation,

Environmental Transport, and Transfer), of which this project is a part of (Figure 1.3). The fifth site was chosen as a radionuclide- and redox-control, which will be discussed in 1.3.3.

1.3.1 The Needle's Eye Natural Analogue Site

Needle's Eye is a site located ~ 17 miles south of the town of Dumfries in south west Scotland, UK (54'889" N, 6'691" W) (Figure 1.3). The site is named after a geological formation (an aperture) ~ 3 m tall by ~ 2 m wide within a section of protruding cliff face. A major vein of exposed pitchblende (UO_2) embedded within the Criffel granodiorite forms a cliff where the intruded rocks are faulted against Carboniferous rocks underlying the Solway Firth estuary (Miller and Taylor, 1966; Milodowski *et al.*, 1990; MacKenzie *et al.*, 1991). This vein was represented the largest and most studied vein found during an initial survey in the area (Miller and Taylor, 1966). The uranium-mineralisation is thought to be 185 ± 20 million years old (Miller and Taylor, 1966; Braithwaite *et al.*, 1997). The site has been well characterised from a geological, hydrological and basic geochemical perspective due to the presence this pitchblende. The upwelling and flow of groundwater through fissures in the fractured granite drains the soil and sediment by the cliff, and interacts with the pitchblende, transporting uranium (U) in the direction of nearby Southwick Merse (salt marsh) and Solway Firth estuary (Milodowski *et al.*, 1990; Hooker, 1991; MacKenzie *et al.*, 1991; Jamet *et al.*, 1993). The flow of groundwater over the pitchblende vein enriches the water with dissolved U and flows in the general direction of the estuary, but before reaching the estuary, becomes retained in a highly organic, humic-rich bog at the foot of the cliff (Braithwate *et al.*, 1997) (Figure 1.4), where as much as 90% of the U is retained in addition to other heavy metals including zinc (Zn) and manganese (Mn) (MacKenzie *et al.*, 1991). This bog was thought to be anoxic, and the accumulation of U in the bog was thought to be due to reduction of the U in the anaerobic conditions of the bog and complexation to indigenous organic matter (Scott *et al.*, 1991; MacKenzie *et al.*, 1991; Jamet *et al.*, 1993; Braithwaite *et al.*, 1997). It is thought as groundwater flows over the pitchblende vein, U from the pitchblende dissolves in the form the uranyl ion (UO_2^{2+}) (U^{6+}) and is transported by the groundwater in the direction of the estuary and be retained in the anoxic bog, where the reduction of U^{6+} to U^{4+} is thought to occur and retain the U (MacKenzie *et al.*, 1991). U^{6+} can form soluble anionic complexes with

naturally occurring ligands (*e.g.*, $[UO_2 (CO_3)_2]^{2-}$), which is the process thought to be occurring in the bog at Needle's Eye (MacKenzie *et al.*, 1991; Jamet *et al.*, 1993; Fuller *et al.*, in prep.). The interaction between U and humic substances (Choppin, 1988) at Needle's Eye have been studied in the relation to the importance of these

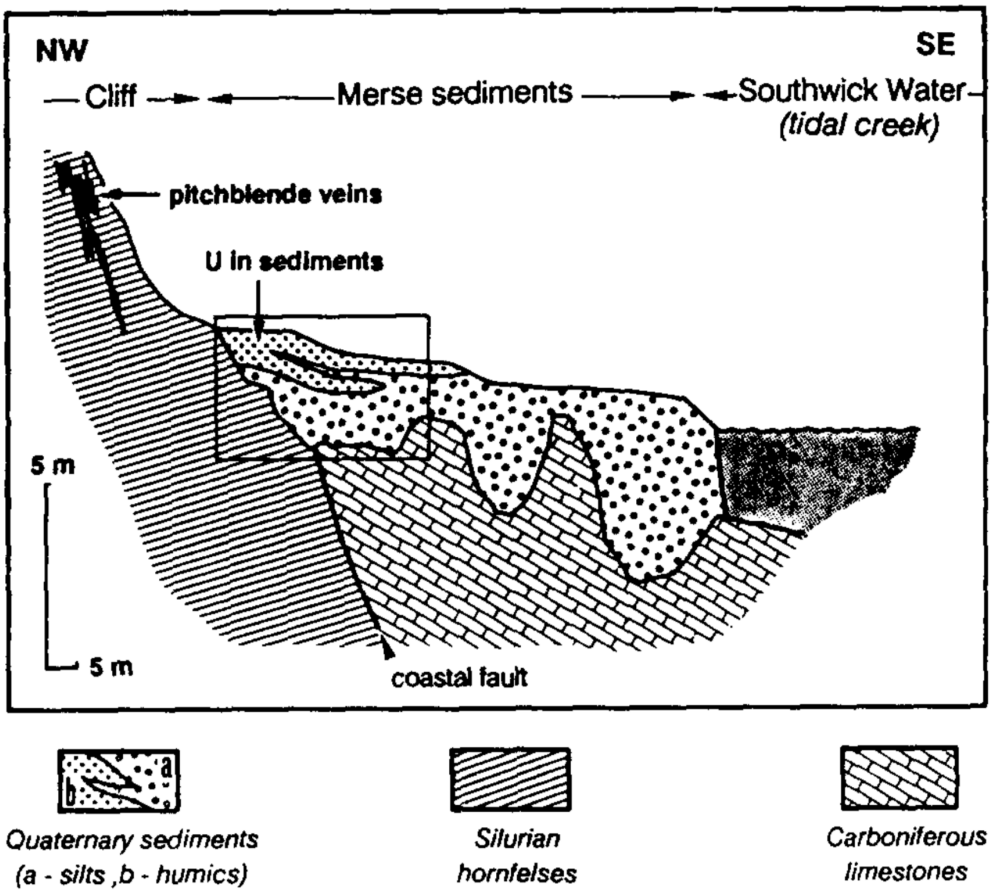


Figure 1.4. From Jamet *et al.* (1993). Uranium mineralisation at Needle's Eye, demonstrating the accumulation of U in humic sediments at the base of the cliff.

interactions on the transport and bioavailability of U, showing that U was bound to lowest molecular weight fractions of humic substances at the site (Zhang *et al.*, 1997). The interaction of U with humic substances can impact the bioavailability of U to the indigenous microbial populations.

The hydrogeology of the site is key to the transport of uranium at the site. Groundwater at the site oxidises the U in the pitchblende from insoluble U^{6+} to soluble U^{4+}, enabling mobility and transport of the U across from the cliff to the bog (MacKenzie *et al.*, 1991). The groundwater flow of the site was also found to have vertical profiles of the site, indicating the upwelling of groundwater from deeper regions of the site (MacKenzie *et al.*, 1991; Jamet *et al.*, 1993) (Figure 1.5), potentially remobilising the reduced U^{4+} in the anoxic bog.

The accumulation of U in this bog is anticipated to be highly localised, and U levels were predicted to be significantly lower elsewhere in the Needle's Eye area and not impacted by U enriched groundwaters, following a site reconnaissance conducted in July 2014, prior to this study. For example, the other regions of the cliff face are not thought to contain similar veins of exposed pitchblende, and therefore other organic-rich bog areas would not accumulate U. This was demonstrated to be the case as part of this project (see Chapter 3). The highly-localised accumulation of U in the bog presents the opportunity to investigate the role and interactions of indigenous microbial communities on elevated U levels in natural environmental conditions.

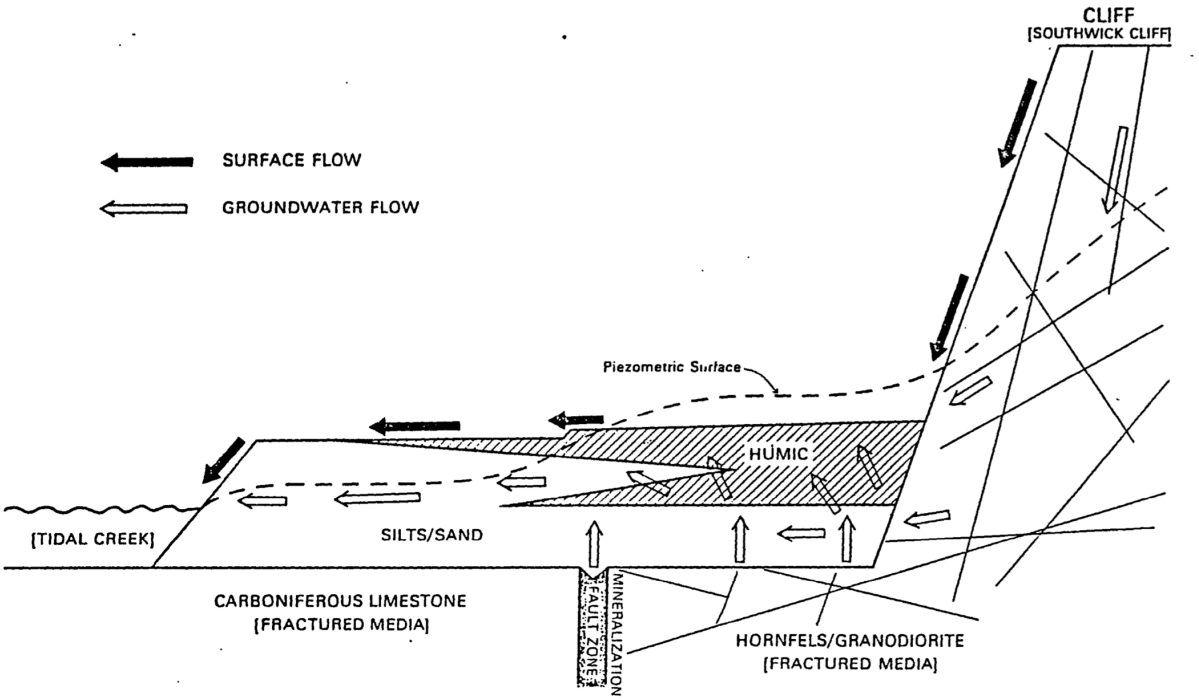


Figure 1.5. From MacKenzie *et al.* (1991). BGS schematic diagram showing the main hydrological features of the Needle's Eye natural analogue site. Pitchblende vein is embedded in the Southwick Cliff.

1.3.2 Sellafield-Derived Radionuclides: The Esk Estuary Saltmarshes and the North East Irish Sea Basin

Two sites, the Esk Estuary saltmarsh by the town of Ravenglass, and the Irish Sea basin, were coupled for the purpose of this study, representing the two sites impacted by Sellafield-derived radionuclides (Figure 1.6). Furthermore, the geophysical and hydrological nature of the two sites has resulted in a hypothesised

source-sink relationship of Sellafield-derived radionuclide deposition at the sites (Kershaw *et al.*, 1990).

Sellafield is the UK's primary nuclear processing and power-generating facility and has been involved in the processing of nuclear fuels since the 1940s initially for weapons manufacturing, later electricity generation, and now in spent-fuel reprocessing and storage. As a result of these activities, many of which encompassed much of the earliest development work using nuclear fuels, there has been a fluctuating but consistent authorised release of radioactive effluents from fission into the local environment since the 1950s, including isotopes of americium (Am), cesium (Cs), neptunium (Np), plutonium (Pu), and U, first atmospherically, and later aqueously (Kershaw *et al.*, 1992; Gray *et al.*, 1995). Maximum rates of discharges occurred in the 1970s due to the UK coal mining strikes leading to increased storage times for Magnox fuels (Gray *et al.*, 1995; MacKenzie *et al.*, 1998), but discharges declined after this period as on-site effluent treatment technologies were developed, *e.g.*, the Enhanced Actinide Removal Plant (EARP) and the Site Ion Exchange Effluent Plant (SIXEP) (Kershaw *et al.*, 1992). Although much lower today, controlled aqueous discharges of radionuclides from Sellafield into the Irish Sea still occur *via* 2.1 km pipes that extend into the eastern Irish Sea (Gray *et al.*, 1995; Muir *et al.*, 2015).

The Irish Sea is a closed macro-tidal, semi-enclosed marine system (Halcrow Group, 2013) connected to the Atlantic Ocean from both the north *via* the North Channel and the south *via* St. George's Channel (Carr and Blackley, 1986). Aqueous discharges of particle-reactive radionuclides from Sellafield released into the Irish Sea become associated with suspended particulates and accumulate in a patch of fine, muddy sediment of the Irish Sea basin known as the Irish Sea mud patches, which has been shown to act as a transient sink for radionuclides discharged from Sellafield (Hetherington and Harvey, 1978; Finegan *et al.*, 2009). However, instead of being deposited at depth by sedimentation as would normally be expected, the radionuclides are reworked in the surficial sediments by indigenous macrofauna, and remobilized into the water column (Kershaw *et al.*, 1988). Reworking as far down as one metre has been observed at the mud patches (Kershaw *et al.*, 1983). These radionuclides are then transported elsewhere by tidal currents, contaminating the coastlines of the UK and Ireland (McCartney *et al.*, 1994; Keogh *et al.*, 2007; Al-Qasmi *et al.*, 2016 and 2017). This is the foundation of the source-sink relationship

between the Irish Sea Mud Patch and the River Esk estuary at Ravenglass (McCartney *et al.*, 1994).

The Ravenglass estuary is ~ 15 km south of the Sellafield site and is comprised of salt marshes covering silt and clay with a bidaily tide (Carr and Blackley, 1986). Here, elevated concentrations of anthropogenic radionuclides have been detected along the intertidal salt marshes that have been brought in from the tide, likely as a result of mixing at the mud patch (Hamilton and Clarke, 1984). The accumulation and sedimentation of Sellafield-derived radionuclides in this region is thought to reflect historical releases from Sellafield (Aston and Stanners, 1981; Morris *et al.*, 2000; Marsden *et al.*, 2006).

At both sites, the biogeochemical cycling, and thus redox activity of Sellafield-derived actinides, has hitherto not been studied. Furthermore, no previous studies have employed contemporary molecular microbial ecological methods to characterise the indigenous microbial communities and functions. A recent study on the same sites by Ray *et al.* (*in prep.*) revealed radionuclide activity profiles in the mud patch that did not appear to reflect Sellafield discharges, but clearly defined profiles at Ravenglass.

Figure 1.6. Map of the UK with inset of the Irish Sea and Cumbrian coastline, and red dots indicating the location of the: Irish sea mud patch; the Sellafield nuclear site, and; the Ravenglass saltmarshes. From Ray *et al.* (in prep.)

1.3.3 The Loch Etive Basins: The 'Control' Site

Loch Etive, which has been studied extensively from a geochemical perspective (see below), is a loch 5 km north of the town of Oban, on the west coast of Scotland and has been well characterised from a hydrography and geochemical perspective. The loch is a glacially overdeepened valley 31.6 km long and between 1.2 – 1.6 km wide, giving a small surface area of ~ 26 km^2 (Wood *et al.*, 1973). It is connected to the North Atlantic Ocean *via* the Firth of Lorne by a narrow sill 300 m wide at Connel (Ridgway and Price, 1987). The loch is comprised of two distinct basins separated by a sill in the middle of the loch at a site named Bonawe (Edwards and Edelsten, 1977). The photic zone, where photosynthetic primary production from diatoms, phytoplankton, and algae is most likely to occur, ranges from 0 – 10 cm of the water column (Wood *et al.*, 1973).

The outer basin that connects to the ocean at Connel runs east-west in direction, is mostly estuarine, and has a maximum depth of 71 m (Figure 1.7). Fresh water from the River Awe enters the outer basin at the western-most point of the loch, at a widening of the basin called Airds Bay. The water in the outer basin is generally well-mixed and oxygenated. Visual assessment during sediment profiling of the outer basin showed a red-brown layer from ~ 0 – 5 cm, below which sediments tended to be grey (Ridgway and Price, 1987). Macrofauna such as worms and bivalves are abundant in the upper sediments, below which macrofaunal activity is considered absent due to anaerobic conditions (Ridgway and Price, 1987). Sedimentation in the outer basin only had a small contribution from phytoplankton, either dead or from faeces (Ansell, 1974). Most organic matter environments similar to the loch is mineralised in the water column and in suspended sediments (De Hass *et al.*, 2002) with only small amounts reaching the seafloor (Libes, 1992). Much of the organic carbon that does reach the floor of the loch is recalcitrant lignin (Coûteaux *et al.*, 1991), which may be used as a biomarker as it is generally resistant to biodegradation (Loh *et al.*, 2008). Sedimentation rates of the outer basin are ~ 0.6 cm per year (Al-Qasmi *et al.*, 2018).

The inner basin runs north-south at an almost-90° angle to the outer basin. The inner basin is much deeper, with a maximum depth of 150 – 170 m, and is generally less mixed and oxygen-poor, and considered to be anaerobic at depth. The River Etive enters the inner basin at the northern-most tip. Loch Etive is

characterised by a lack of interaction between these two basins due to the shallow sill at Bonawe, except for periodic flushing of the inner basin with fresher, oxygenated waters, caused by overflowing of the sill due to spring tides, regenerating the deeper basin water around every 16 months (Edwards and Edelsten, 1977; Ridgway and Price, 1987). Sedimentation rates of the inner basin are slower than that of the outer basin, at ~ 0.3 cm per year (Al-Qasmi *et al.*, 2018).

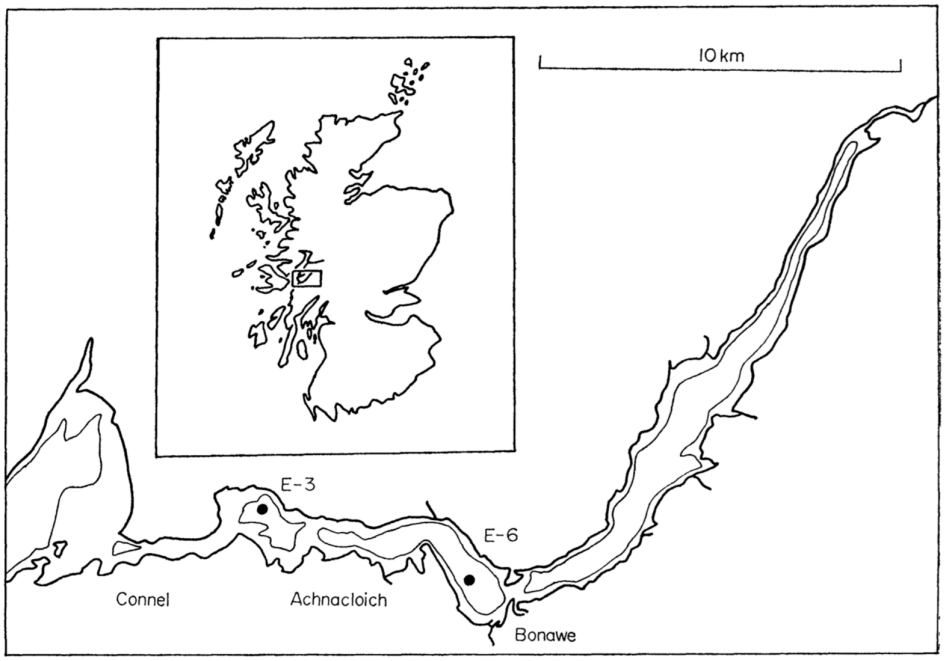


Figure 1.7. Map of Loch Etive with inset of map of Scotland. Contours denote 20 cm depth. Connel is the opening to sea. Bonawe is the sill between the shallow and deep basin. From Wood *et al.* (1973).

The presence of heavy metal contaminants, likely from an anthropogenic source, has been identified in sediments collected from Loch Etive (Ridgway and Price, 1987). These metals, including Zn, Pb, and ^{137}Cs of which the later probably originated from the Sellafield nuclear plant in Cumbria (England), were found to be enriched in sediments from Loch Etive, with higher values found in the deeper, inner basin, likely due to less biomixing. In the context of the subject focus of this thesis, a recent study showed that a range of natural ($^{234, 238}$U) and artificial (^{137}Cs, ^{241}Am, ^{238}Pu, $^{239/240}$Pu, and ^{236}U) radionuclides from Sellafield have been deposited in the sediments of Loch Etive (Al-Qasmi *et al.*, 2018). This study found that ^{137}Cs and ^{236}U abundance profiles in the sediments reflected historical releases from Sellafield during the 1970s, while Pu and Am fraction profiles did not, and that ^{236}U was not affected by redox conditions in the outer basin, *i.e.*, was not reduced at depth in the Loch, indicating the lack of sufficiently reducing conditions. The authors attribute the

uniform sedimentation profiles of Pu and Am in the Loch to the Irish Sea Mud Patch (see Chapter 4) being a constant source of radionuclide contamination to the Loch.

The contrasting conditions of the two basins, despite being of the same loch, and the well-characterised hydrography and geochemistry of the site, made it an ideal location for a natural laboratory to investigate radionuclide deposition from Sellafield at a greater distance than Ravenglass. However, the microbial community of the loch has, hitherto, not been investigated using contemporary molecular biological or high-throughput DNA sequencing techniques. The site was initially chosen as a 'control' site for the entire project, as it was believed to be a site that had negligible levels of redox-active radionuclide contamination (*i.e.*, not greater than background levels) from either geogenic or anthropogenic sources, and additionally, as a sediment system with redox conditions that followed the normally-expected or 'classical', thermodynamically-controlled vertical stratification of microbial redox processes with depth (Thauer *et al.*, 1977), due to the constant sedimentation and lack of mixing observed in the inner basin. In light of the publication of the radionuclide study by Al-Qasmi *et al.* (2018), the focus of this particular study shifted from that of treating the loch as a radionuclide control, to one of being a redox control – where the expected transitions of microbial community taxonomic and functional assemblages with increasing depth and related redox succession could be directly compared with that observed at other sites in the project.

1.4 The Advent of High Throughput DNA Sequencing

The evolution of modern molecular microbial ecology methods has been focused on the advent of DNA and RNA amplification and sequencing methods. Extracting and targeting the DNA and RNA of microbial populations directly from soil and sediment samples removes the need for culture-dependent methods. Culture-dependent methods suffer from the fact that for unknown reasons, only between 0.1 – 10% of known bacteria and archaea are culturable (Staley and Konopa, 1995; Head *et al.*, 1998), while targeting DNA and RNA in theory captures the entire representative microbial community of a sample (Daniel, 2005). DNA molecular methods have themselves evolved over the last few decades, culminating in the contemporary highly-parallelised, high-throughput DNA sequencing methods that are widely used to study microbial ecology today, which generate many millions of sequences from diverse environmental samples in hours (Scholz *et al.*, 2012). Two

main approaches to using DNA sequencing to investigate microbial ecology will be discussed here that were used during this project: The targeted amplification and sequencing of target marker gene DNA sequences, and; untargeted shotgun metagenome sequencing.

1.4.1 Marker Gene DNA Sequencing Methods

Polymerase chain reaction (PCR) is performed using primers to amplify specific target marker genes of genomic DNA extracted from soils and sediments. This, under ideal conditions and efficiency, results in a doubling of the amount of target DNA copies per PCR cycle, which over the 20 – 40 cycles of PCR commonly performed, results in many millions of copies and significantly increasing the amount of target DNA over background genomic DNA (Malik *et al.*, 2008). These amplicons then undergo a series of clean up steps to remove excess PCR reagents and non-target DNA and then sequenced (see Chapter 2.5).

The identification of bacterial and archaeal populations in habitats has coalesced around the targeting and sequencing of the marker 16S ribosomal RNA (rRNA) gene (Head *et al.*, 1998; Muyzer, 1999). The 16S rRNA gene, as characterised by Woese (1987), contains many properties that make it an ideal marker gene target for the characterisation of microbial populations, including: highly conserved sequence domains that are targetable by primers; variable regions interspersed between conserved regions, that allow for phylogenetic and taxonomic assignment; the large amount of rRNA in cells; not affected by lateral gene transfer; a long length of ~ 1,500 base pairs (bp) that provides sufficient sequence information, and; the curation of rRNA reference databases for comparative sequence analysis (McCann, 2012; Amann and Ludwig, 2000; Case *et al.*, 2007; Gutell *et al.*, 1994; Head *et al.*, 1998; Olsen *et al.*, 1986; van de Peer *et al.*, 1996). Comparative analysis of 16S rRNA gene sequencing data enables the ecology, diversity, and evolution of microbial diversity of a sample (Head *et al.*, 1998).

By comparison, the identification of fungal populations in soils and sediments focuses on the targeting and sequencing of the internal transcribed spacer (ITS) region, a region that has been used as a universal marker gene for identifying fungal taxonomy in microbial ecology for many years, with many primers able to achieve high coverage of the fungal kingdom (White *et al.*, 1990; Gardes and Brun *et al.*,

1993). The ITS region is found in-between the small 18S subunit and the large 28S subunit of the rRNA gene. It is variable in length, usually around 600 bp, but is a hypervariable region, commonly ranging between 500 – 750 bp in length (Op de Beeck *et al.*, 2014). The ITS region is comprised of three parts: the ITS1 region; the 5.8S subunit, and; the ITS2 region (Toju *et al.*, 2012). However, as the ITS region is hypervariable in its length, the alignment of sequenced reads is very difficult due to the lack of conserved regions or the inability to reliably identify conserved regions in hypervariable lengths between different fungal taxa, impeding phylogenetic analysis of ITS sequenced reads (Toju *et al.*, 2012; Schoch *et al.*, 2012). Despite efforts to identify an alternative marker gene region, no such region has yet been identified with suitable sequence conservation for primer targeting, with sufficient variability for taxonomic assignment and a relative homogeneity of total region length (Schoch *et al.*, 2012).

Using universal marker genes to target microbial community assemblages has allowed for the development and scaling of efficient and well-understood techniques. Additionally, the curation of reference databases for microbial phylogeny and taxonomy enables quick identification and classification of sequenced reads. Currently, the latest version of the UNITE reference database of ITS sequences contained 8,997 reference sequences (UNITE version 7.2, Kõljalg *et al.*, 2013), while the latest version of SILVA for 16S contained 695,171 reference sequences (version 132, Quast *et al.*, 2013).

The methods used for DNA marker gene sequencing in this project are described in Chapter 2.5.

1.4.2 Marker Gene Sequencing Bioinformatics Pipeline Analysis

Software tools such as QIIME (Caporaso *et al.*, 2010) and mothur (Schloss *et al.*, 2009) have emerged to aid in the analysis of multiplexed and large marker gene sequencing datasets being generated by high-throughput sequencing. These tools aim to standardise and simplify the pipeline workflows of marker gene sequencing analysis. The tools differ in their precise methods, but generally process raw sequencing data by: quality control, by removing low-quality, short, and chimeric (amalgamation of multiple reads) sequences; identify operational taxonomic units (OTUs), or amplicon sequence variants (ASVs); classify taxonomy; align reads and

create phylogenetic trees, and; perform a range of alpha and beta diversity analyses to describe community structure and assemblages between samples. The main difference between QIIME and Mothur is that QIIME is written in the Python programming language, and thus is compatible with only the UNIX kernel, while Mothur is written in C and compatible with any kernel. There are also differences in the defaults used in each step, that produce slightly different results if the user is not aware of these defaults (chiefly, the method used for creating OTUs). Both rely on reference databases to classify taxonomy and to align reads *e.g.*, Greengenes, RDP, SILVA.

A recent development in the field of marker gene sequencing bioinformatics is the emergence of tools which aim to create functional gene abundance profiles using the phylogeny and taxonomy derived from marker gene sequencing, to create a gene abundance output analogous to a metagenome profile. Since 2013, several tools have been published and include PICRUSt (Langille *et al.*, 2013), Tax4Fun (Asshauer *et al.*, 2015), Piphillin (Iwai *et al.*, 2016), and PAPRICA (Bowman and Ducklow, 2015). Though each tool differs in its method, the goal of the tools is to link OTUs/ASVs from marker gene analysis to their closest fully sequenced relative, and create a predicted genome for that OTU, across all OTUs, and all can output genes in the KEGG Ortholog format. The main principle of most of the tools is to use hidden state prediction (*via* Hidden Markov Models) to infer gene content of an OTU based on linking the 16S rRNA gene sequence to a fully sequenced relative. While any algorithm- and prediction-based tool has the potential to produce erroneous results, the accuracy of these tools can be directly tested and compared against real metagenome data. Furthermore, as the taxonomic and genomic reference databases increase constantly, the reliability of these inference tools will also increase, as the distance between an OTU and a closest relative sequence decreases. Using gene abundance inference tools allows the deep 16S rRNA gene sequencing data to do ‘double-duty’, and potentially skip the need for an equivalently deep (and thus very expensive) metagenome sequencing.

PICRUSt (Langille *et al.*, 2013) uses ancestral state reconstruction to link an OTU to its closest fully sequenced ancestor and infers the genome of the OTU based on the phylogenetic distance of the OTU to the whole-genome ancestor. It uses the Greengenes (DeSantis *et al.*, 2006) reference taxonomy database to create a reference OTU phylogenetic tree, and links this to a gene content table from a

reference genome database (*e.g.*, KEGG) to create a predicted gene content for each OTU in the Greengenes database. This is then used to infer the metagenome of the sample dataset by combining the resulting gene content predictions for all the OTUs based on the relative abundances of the 16S rRNA genes and corrects for expected 16S rRNA gene copy number. The output is a table of genes families and abundances of the community of a sample. Because an OTU within a tree always has a neighbour, PICRUSt can link any OTU, even over large phylogenetic distances. This means that microbial phyla that are poorly represented in the reference genome database (*e.g.*, KEGG) can potentially introduce a lot of uncertainty in the predicted genomic profile of those OTUs. Additionally, the Greengenes reference database has not been updated since August 2013, over four years at time of writing.

Tax4Fun (Asshauer *et al.*, 2015) uses a precomputed association matrix and a precomputed metabolic reference profile to create an inferred functional gene profile, to link 16S rRNA genes with the functional profile of a sequenced genome, by using nearest-neighbour identification based on a minimum 16S rRNA sequence similarity. Tax4Fun uses the SILVA-labelled OTU abundances to link OTUs with reference 16S rRNA sequences in the SILVA database. The SILVA-linked 16S rRNA profile is then transformed to a relevant taxonomic profile from KEGG organisms (*via* a precomputed association matrix). Next, the estimated abundances of KEGG organisms are normalised by the 16S rRNA copy number (derived from the NCBI genome annotations). Then, the normalised taxonomic abundances are used to linearly combine the precomputed functional profiles of the KEGG organisms for the prediction of the functional profile of the microbial community. Since Tax4Fun requires the precomputation of the association matrix and metabolic reference profiles, the output is limited to the versions of SILVA and KEGG used to create these precomputed files. As a result, Tax4Fun is using older versions of both databases – KEGG 64.0 from October 2012 and SILVA 119 from July 2014. These two databases are now onto versions 84.0 and 132 respectively. KEGG 84.0 contains around 4,000 extra genes than KEGG 64 (a 25% increase), while SILVA 132 contains over 150,000 extra non-redundant reads (a 28% increase).

Piphillin (Iwai *et al.*, 2016) claims to have a superior method of gene abundance inference because it does not rely on using certain (outdated) reference databases to create precomputed profiles and can be easily updated regularly.

Piphillin does not use a phylogenetic tree, and instead simply links an OTU to the closest fully sequenced relative by matching the 16S rRNA gene sequence to the closest match in the genome database. If an OTU matches with two or more fully sequenced genomes, the genomes are split equally to infer the OTU genome. However, the authors state that while they found Piphillin to be superior at predicting the metagenome of human microbiome samples, Piphillin was deemed inferior to Tax4Fun on environmental samples.

PAPRICA (Bowman *et al.*, 2015) takes a similar approach to PICRUSt and Tax4Fun by relying on phylogenetic trees. Additionally, it does not use an OTU-based approach, instead relying on unclustered 16S rRNA reads, aiming to increase the resolution between species that have a higher similarity than the 97% usually ascribed.

Tax4Fun was used throughout this project as it is both open-source and uses the more regularly updated SILVA reference database.

The above gene abundance inference tools all (for the exception of PAPRICA) have one step in common: the clustering of 16S rRNA sequences into OTUs. Various methods and options exist with regards to OTU picking steps. In QIIME, two main approaches exist, open- and closed-reference OTU picking, and *de novo* OTU picking. Briefly, open- and closed-reference OTU picking search the 16S rRNA reads against a reference database (*e.g.*, Greengenes, SILVA) using a similarity threshold (*e.g.*, 97% similarity, 99% similarity between sequenced reads), and all reads that match against the reference database are considered OTUs. Reads that failed to match are then clustered amongst themselves at the given similarity threshold. These failed clusters are then searched again against the reference database, where a number will now have matches in the database and become OTUs. The failures from this step are then considered *de novo* OTUs in an open-reference table or omitted in a closed-reference table. *De novo* OTU picking simply clusters all 16S rRNA reads together at a given similarity threshold, without using a reference database. Taxonomy is assigned in all methods from the reference database. Numerous different tools exist to perform OTU picking within QIIME, including uclust, usearch, vsearch, sortmerna, and swarm (amongst many others). The 97% similarity threshold has been discussed in great detail at this stage, and it is widely accepted that 97% similarity is too low for many prokaryotic species (Callahan *et al.*, 2017). Another widely-discussed issue is with the inflation of *de novo* OTUs in QIIME1,

which characterised by the OTU picking methods massively overestimating the number of *de novo* OTUs in a sample, which has a detrimental impact on alpha diversity indices and relative abundances of the other OTUs (Edgar, 2017).

As such, in QIIME2, 16S rRNA reads are no longer clustered at a similarity threshold against a reference database to create OTUs, and instead sequences are clustered together at 100% similarity (not against a reference database) and a representative sequence is chosen to infer an individual read, known as amplicon sequence variants (ASVs) using the DADA2 tool (Callahan *et al.*, 2017; Callahan *et al.*, 2016). These ASVs are distinguished from one another at the single nucleotide level (*i.e.*, 100% similarity), using a range of algorithms that perform parametric error-rate modelling to identify mis-sequenced bases, and infer which sequences represent statistically different ASVs (Callahan *et al.*, 2016). These ASVs are treated as unique biological units that represent true biological diversity between sequences that likely represent evolutionary events (Callahan *et al.*, 2016; Callahan *et al.*, 2017). This approach removes the reliance of reference databases and any biases that approach introduced, *i.e.*, taxonomic clades that are under-represented in the reference database may be classed as *de novo* OTUs and excluded from any closed-reference analysis. Taxonomy is then classified independently to the ASV using a naïve-Bayes trained classifier as a completely separate step, using extracted reads of the target sequence region from a reference database (Bokulich *et al.*, 2017). This mitigates any potential inflation of OTU/ASV numbers. The different numbers of OTUs and ASVs created by a range of picking methods from QIIME1 and QIIME2 is discussed in Chapter 2.6.1. A full description of a typical marker gene sequencing pipeline is in Chapter 2.6.

1.4.3 Shotgun Metagenome Sequencing Methods

As opposed to targeting and amplifying specific marker genes to sequence taxonomic and phylogenetic assemblages of microbial populations, an alternative approach is to sequence a random, untargeted, selection of the microbial genomes in DNA extracted from soils and sediments. Shotgun sequencing can be used to profile not only the taxonomic profiles of microbial communities due to sequencing of the 16S rRNA and ITS genes, but of the functional capacity of a microbial community by the sequencing of sections of genomes associated with a sample (Shah *et al.*, 2011; Quince *et al.*, 2017). However, due to its untargeted nature, shotgun metagenome

sequencing does not reveal microbial community assemblages to the same depth as marker gene sequencing, especially in complex environments (Shah *et al.*, 2011). Additionally, the presence of functional genes in a metagenome sequencing dataset does not necessarily imply that the microorganisms were performing that function at time of sampling, merely that they possessed the genomic capacity of that function. The MG-RAST (Meyer *et al.*, 2008) platform is a commonly-used online pipeline that produces a wide range of phylogenetic, taxonomic, and functional annotation data for shotgun metagenomic sequencing data. It is also possible to manually analyse data, A basic pipeline for analysing metagenome NGS data is: Assembling overlapping reads into contigs, using a variety of algorithms *e.g.*, greedy, overlap-layout-consensus, or de Bruijn graph methods; placing these contigs into groups, intended to reflect individual species/genomes ('binning'); if a bin has sufficient number of contigs, create a scaffold genome, *i.e.*, a reconstructed draft genome of one species; then annotate all the assembled contigs and bins to derive function of the genes (Scholz *et al.*, 2011; Sedlar *et al.*, 2017; Quince *et al.*, 2017). A profile of all annotated genes and their abundances can be produced. However, many different programs exist that perform many of same steps, and do so differently, with each claiming to be superior to the other. This can make it difficult to know which one to use and when. Many review articles exist that compare various methods for the same step (Thomas *et al.*, 2012; Sedlar *et al.*, 2017; Ghurye *et al.*, 2016; Breitwieser *et al.*, 2017).

1.4.4 Other High-Throughput Techniques

In addition to DNA-based molecular high-throughput sequencing techniques, alternative methods to characterise the taxonomic and functional assemblages of indigenous microbial communities exist.

Metatranscriptomics, commonly referred to as RNA-seq, is the method of sequencing targeting messenger RNA (mRNA) fragments in a system in lieu of genomic DNA (gDNA). While many of the sequencing methods, platforms, and data analysis are shared between RNA and DNA sequencing, the sample preparation and scientific value differ. The direct sequencing of mRNA of a system enables a snapshot of the metabolic processes that were actively being performed by the microbial community at time of sampling – in contrast to the longer-term functional potential that is acquired by metagenomics. Briefly, mRNA that has been transcribed

in order to be translated into proteins and enzymes by the microorganisms are converted into complementary DNA (cDNA) and sequenced using a high-throughput sequencer. Sequenced reads are assembled *i.e.*, concatenated into much longer contiguous reads known as contigs, which are annotated with a genomic function using a reference database similar to metagenome data. Contigs can be aligned against a reference genome to derive transcript counts. Counts of transcripts are used to infer the presence and abundance of the related genomic function, and differential abundance of transcripts calculated to derive functional differences between samples. The primary advantage of RNA-seq and sequencing the metatranscriptome of a microbial community is that it enables a snapshot view of the genes being expressed, and thus the proteins and enzymes directly being synthesised by the microbial community, at time of extraction, which directly reflects microbial response to the surrounding environmental conditions. This is arguably superior to marker gene or metagenomic sequencing which only provide taxonomic assemblage and genomic functional potential respectively. However, the method is not without its drawbacks. Firstly, mRNA is short-lived, with *in-vitro E. coli* mRNA shown to have a half-life of 6.8 minutes (Selinger *et al*., 2003). This can make extracting mRNA from complex and hard-to-reach environments difficult and requires snap-freezing with liquid nitrogen. Secondly, less than 5% of the total RNA in soil is mRNA, with the remaining 95% being almost entirely comprised of ribosomal RNA (Neidhardt *et al*., 1996; Sharma *et al*., 2012). Furthermore, prokaryotic mRNA generally lacks the poly-adenosine-monophosphate tail (polyadenylation, also referred to as poly(A) tail) on the 3`-end, making it more difficult to isolate mRNA from rRNA (McGrath *et al*., 2008). Lastly, the relative cost of RNA-seq compared to marker gene sequencing, and to a lesser extent metagenomic sequencing, is significantly highly, predicated on the fact that a much higher sequencing depth is required in order to obtain sufficient coverage of the transcriptome and derive an adequate analysis of the metabolic function of the microbial community.

Likely due to the prohibitive methods and costs associated with RNA-seq targeting prokaryotes in natural environments, examples of RNA-seq being applied in the field of environmental microbial ecology are relatively rare. Recent publications include Holmes *et al*. (2017) which used RNA-seq to investigate electron transfer between potentially syntrophic methanogenic archaea and metal-reducing *Geobacter* spp. as inferred by the abundance of transcripts coding for electrically conductive pili

genes. Xia *et al*. (2014) used RNA-seq in tandem with metagenomic sequencing to characterise cellulose degradation and associated methanogenesis.

Metaproteomics is the method of analysing the proteins synthesised by a population. Proteomics broadly reflects the transcriptome, and therefore is analogous to metatranscriptomics in that it characterises the metabolic function of population, and that the mRNA sequenced by RNA-seq is inevitably used to produce proteins. Modern metaproteomic methods are centred around variations of mass spectrometry (MS) which measure the mass of proteins in a method called peptide mass fingerprinting, where the observed mass of a protein (the peak) is matched to a protein in a reference database (Bastida *et al*., 2014). Soil proteomics is susceptible to interference by contaminants such as humic substances, similar to DNA and RNA sequencing (Bastida *et al*., 2014). With regards to microbial ecology of natural environments, metaproteomics has generally been less favoured than DNA or RNA sequencing techniques. Benndorf *et al*. (2007) used proteomics to characterise proteins in contaminated soil and groundwater, identifying proteins expected in the systems, while Williams *et al*. (2010) used metaproteomics in tandem with 16S rRNA marker gene sequencing to characterise proteins and thus microbial function in soils amended with the hydrocarbon toluene.

Metabolomics is the method of analysing metabolites produced by organisms and microorganisms within a system during normal cell metabolism (Tang, 2011). Metabolites reflect the genome, transcriptome, and proteome of cell, and describe the physiological state of the cell (Markley *et al*., 2017). MS and nuclear magnetic resonance (NMR) are the main methods used for metabolomics, where peaks are matched to compounds in a reference database. Tremaroli *et al*. (2009) used metabolomics to investigate the response in cultures of the soil commensal *Pseudomonas pseudoalcaligenes* to the toxic metalloid tellurite, where NRM-based metabolomics showed metabolic changes in bacterial cultures in response to the addition of tellurite.

1.4.5 Geochemical Analysis Methods

A range of geochemical analytical methods are available that aim to determine chemical composition, concentration, and redox speciation of compounds either dissolved in solution or precipitated in bulk soils and sediments. These tools are often

used in tandem with one another to encapsulate both soluble and insoluble redox species of elements of interest within a sample. The presence and valence of key redox-active compounds within an environmental sample can be directly attributed to the presence of microbially-mediated metabolic processes within that environment. A brief description of the methods used in this study follows.

Inductively coupled plasma mass spectrometry (IMP-MS) detects a range of metals and non-metals that are in solution. Solution is pumped into the main chamber and argon plasma, generating positively charged ions. Mass spectrometry is used to characterise the ions based on their mass-to-charge ratio. The strength of the signal on the MS output pertains to ion concentration. Standards are used to calibrate the MS. ICP-MS is able to detect elements of an atomic weight up until uranium (Thomas, 1993). The detection limit of ICP-MS is from parts per billion (ppb) (Lloyd *et al.*, 2007) to parts per trillion (ppt) (Thomas, 1993). Fe and Mn are soluble in their reduced 2+ state, while U is soluble when oxidised as U^{6+}, so detecting these metals *via* ICP-MS within a sample can be used to directly infer the presence of microbially-mediated electron transfer to these elements, *i.e.*, the detection of Fe or Mn *via* ICP-MS in a sample indicates the presence of the reduced species of these elements, and therefore potential for bioreduction.

X-ray fluorescence (XRF) is the method of determining the major and trace elemental composition of both solid and aqueous samples. X-rays are directed at the sample, exciting the electrons of an atom from an inner orbit to an outer orbital shell, and as that electron returns to its inner shell, fluorescent X-rays are released and detected. The measurement of the energy of the released X-ray is used to determine the element hit by the emitted X-ray, and intensity of the wavelength is used to determine concentration. As it targets more than just metals, it has a wider range of elemental targets than ICP-MS, as well as being able to detect elements in solid as well as aqueous form (Rollinson, 1993). Pressed powder discs of bulk soil or sediment can be used to investigate the presence of redox-active compounds, and on the contrary to aqueous-based ICP-MS (Rollinson, 1993), the detection of oxidised Fe^{3+} and Mn^{4+} within a sample implies a lack of bioreduction, whereas detection of reduced U^{4+} *via* XRF would imply bioreduction.

X-ray absorption spectroscopy (XAS) is a synchrotron-based method of directly elucidating the oxidation state of metals (Lloyd *et al.*, 2007). XAS is itself split into two main methods. X-ray absorption near-edge structure (XANES) is used to

characterise bulk minerals (Henderson *et al*., 2014) and derive oxidation state in aqueous solution (Lloyd *et al*., 2007) and can be used to determine U oxidation state (Kemner and Kelly, 2007). Extended X-ray absorption finer structure (EXAFS) can be used to determine the coordination of an atom of interest and derive information of the surrounding chemical environment (Lloyd *et al*., 2007; Kemner and Kelly, 2007).

1.5 Research Aim and Objectives

This research aims to use a wide range of microbial molecular techniques to create profiles of microbial diversity and catabolic function of a range of radionuclide contaminated soils and sediments, in an attempt to characterise the prevailing biogeochemical conditions of the sites and then predict the long-term fate of radionuclide contaminants. To the author's knowledge, this will be the first time such a range of methods have been employed in this fashion, specifically with regards to the aim of inferring biogeochemical cycling and redox conditions from microbial profiles, including the use of microbial functional profile prediction tools, to make assumptions of the redox-fate of actinide contaminants in the sites.

Furthermore, specific inferences of geochemical conditions and radionuclide speciation will be compared with directly measured geochemical analysis using a range of techniques, including ICP-MS, XRF, and HERFD-XANES. This will be the first time that microbial community taxonomic and functional assemblages *in situ* at a natural environment will be used to directly infer prevailing geochemical conditions, and those inferences compared with directly measured geochemical data.

1.6 References

Akob DM, Mills HJ, Kostka JE (2006) Metabolically active microbial communities in uranium-contaminated subsurface sediments. FEMS Microbiology Ecology, 59: 95–107

Al-Qasmi H, Law GTW, Fi LK, Livens FR (2016) Origin of artificial radionuclides in soil and sediment from North Wales. Journal of Environmental Radioactivity, 151(1): 244–249

Al-Qasmi H, Law GTW, Fifield KL, Howe J, Brand T, Cowie GL, Law KA, Livens FR (2017) Deposition of Artificial Radionuclides in Sediments of Loch Etive, Scotland. Journal of Environmental Activity, 187: 45–52

Al-Qasmi H, Law GTW, Fifield LK, Howe JA, Brand T, Cowie GL, Law KA, Livens FR (2018) Deposition of artificial radionuclides in sediments of Loch Etive, Scotland. Journal of Environmental Radioactivity, 187: 45–52

Amann R, Ludwig W (2000) Ribosomal RNA-targeted nucleic acid probes for studies in microbial ecology. FEMS Microbiology Reviews, 24(5): 555-565

Anderson RT, Vrionis HA, Ortiz-Bernad I, Resch CT, Long PE, Dayvault R, Karp K, Marutzky S, Metzler DR, Peacock A, White DC, Lowe M, Lovley DR (2003) Stimulating the In Situ Activity of *Geobacter* Species To Remove Uranium from the Groundwater of a Uranium-Contaminated Aquifer. Applied and Environmental Microbiology, 69(10): 5884–5891

Ansell AD (1974) Sedimentation of Organic Detritus in Lochs Etive and Creran, Argyll, Scotland. Marine Biology, 27: 263–273

Asshauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomics 16S rRNA data. Bioinformatics, 31(7): 2882–2884

Aston S, Stanners DA (1981) Americium in intertidal sediments from the coastal environs of Windscale. Marine Pollution Bulletin, 12(5): 149–153

Berner-Lattimi R, Veeramani H, Vecchia ED, Junier P, Lezama-Pacheco JS, Suvorova EI, Sharp JO, Wigginton NS, Bargar JR (2010) Non-uraninite products of microbial U(VI) reduction. Environmental Science and Technology, 44(24): 237–247

Bastida F, Hernández T, García C (2014) Metaproteomics of soils from semiarid environment: Functional and phylogenetic information obtained with different protein extraction methods. Journal of Proteomics, 101: 31–42

Belli KM, DiChrisitana TJ, van Cappellen P, Taillefert M (2015) Effects of aqueous uranyl speciation on the kinetics of microbial uranium reduction. Geochimica et Cosmochimica Acta, 157: 109–124

Benndorf D, Balcke GU, Harms H, von Bergen M (2007) Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. The ISME Journal, 1: 224–234

Berner RA (1980) Early Diagenesis: A Theoretical Approach. Princeton University Press, New Jersey

Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso JG (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome, 6(90): https://doi.org/10.1186/s40168-018-0470-z

Bostick BC, Vairavamurthy MA, Karthikeyan KG, Chorover J (2002) Cesium Adsorption on Clay Minerals: An EXAFS Spectroscopic Investigation. Environmental Science and Technology, 36(12): 2670–2676

Boukhalfa H, Icopini GA, Reilly SD, Neu MP (2007) Plutonium(IV) Reduction by the Metal-Reducing Bacteria *Geobacter metallireducens* GS15 and *Shewanella oneidensis* MR1. Applied and Environmental Microbiology, 73(18): 5897–5903

Bowman JS, Ducklow HW (2015) Microbial Communities Can Be Described by Metabolic Structure: A General Framework and Application to a Seasonally Variable, Depth-Stratified Microbial Community from the Coastal West Antarctic Peninsula. PLoS One, 10(8): e0135868. doi:10.1371/ journal.pone.0135868

Braithwaite A, Livens FR, Richardson S, Howe MT, Goulding KWT (1997) Kinetically controlled release of uranium from soils. European Journal of Soil Science, 48: 661–673

Breitwieser FP, Lu J, Salzberg SL (2017) A review of methods and databases for metagenomic classification and assembly. Briefings in Bioinformatics, bbx120, https://doi.org/10.1093/bib/bbx120

Brookshaw DR, Pattrick RAD, Lloyd JR, Vaughan DJ (2012) Microbial effects on mineral-radionuclide interactions and radionuclide solid-phase capture processes. Mineralogical Magazine, 76(3): 777–806

Burgos WD, McDonough JT, Senko JM, Zhang G, Dohnalkova AC, Kelly SD, Gorby Y, Kemner KM (2008) Characterization of uraninite nanoparticles produced by *Shewanella oneidensis* MR-1. Geochimica et Cosmochimica Acta, 72(20): 4901–4915

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 13: 581-583

Callahan Bj, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. The ISME Journal, 11: 2639–2643

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Pena A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7(5): 335-336

Cardenas E, Wu W-M, Leigh MB, Carley J, Carroll S, Gentry T, Luo J, Watson D, Gu B, Ginder-Vogel M, Kitanidis PK, Jardine PM, Zhou J, Criddle CS, Marsh TL, Tiedje JM (2010) Significant Association between Sulfate-Reducing Bacteria and Uranium-Reducing Microbial Communities as Revealed by a Combined Massively Parallel Sequencing-Indicator Species Approach. Applied and Environmental Microbiology, 76(20): 6778–6786

Carr A, Blackley MWL (1986) Implications of Sedimentological and Hydrological Processes on the Distribution of Radionuclides the Example of a Salt Marsh near Ravenglass: 529–543

Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, Kjelleberg S (2007) Use of 16S rRNA and rpoB Genes as Molecular Markers for Microbial Ecology Studies. Applied and Environmental Microbiology, 73(1): 278-288

Cerrato JM, Ashner MN, Alessi DS, Lezama-Pacheco JS, Beriner-Latmani R, Bargar JR, Giammar DE (2013) Relative reactivity of biogenic and chemogenic uraninite and biogenic noncrystalline U(IV). Environmental Science and Technology, 47(17): 9756–9763

Chiang PN, Wang MK, Huang PM, Wang JJ, Chiu CY (2010) Cesium and strontium sorption by selected tropical and subtropical soils around nuclear facilities. Journal of Environmental Radioactivity, 101(6): 472–481

Choppin GR (1988) Humics and Radionuclide Migration. Radiochimica Acta, 44/45: 23–28

Cologgi DL, Lampa-Pastirk S, Speers AM, Kelly SD, Reguera G (2011) Extracellular reduction of uranium via *Geobacter* conductive pili as a protective cellular mechanism. Proceedings of the National Academy of Sciences, 201108616: DOI:10.1073/pnas.1108616108

Coûteaux M-M, Mousseau M, Célérier M-L, Bottner P (1991) Increased atmosphereic CO_2 and litter quality: decomposition of sweet chesnut leaf litter with animal food webs of different complexitities. OIKOS, 61: 54–64

Daniel R (2005) The metagenomics of soil. Nature Reviews Microbiology, 3(6): 470-478

de Hass H, van Weering TCE, de Stiger H (2002) Organic carbon in shelf seas: sinks or sources, processes and products. Continental Shelf Research, 22: 691–717

DeSantis TZ, Hugenholtz P, Rojas LM, Keller BK, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Applied and Environmental Microbiology, 72(7): 5069–5072

Du X, Boonchayaanant B, Wu W-M, Fendorf S, Bargar J, Criddle CS (2011) Reduction of Uranium(VI) by Soluble Iron(II) Conforms with Thermodynamic Predictions. Environmental Science and Tecnology, 45: 4718–4725

Edgar RC (2017) Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. PeerJ, 5: e3889. https://doi.org10.7717/peerj.3889

Edwards A, Edelsten DJ (1977) Deep Water Renewal of Loch Etive: A Three Basin Scottish Fjord. Estuarine and Coastal Marine Science, 5: 575–595

Ewing RC (1999) Nuclear waste forms in actinides. PNAS, 96: 3432–3439

Ewing RC, Runde W, Albrech-Schmitt TE (2010) Environmental impact of the nuclear fuel cycle: Fate of actinides. MRS Bulletin: 35(11): 859–866

Francis AJ, Dodge CJ, Lu F, Halada GP, Clayton CR (1994) XPS and XANES Studies of Uranium Reduction by Clostridium sp. Environmental Science and Technology, 28(4): 636–639

Francis AJ (1998) Biotransformations of uranium and other actinides in radioactive wastes. Journal of Alloys and Compounds, 271: 78–84

Francis AJ (2007) Microbial mobilization and immobilization of plutonium. Journal of Alloys and Compounds, 445-445: 500–505

Fike DA, Bradley AS, Rose CV (2015) Rethinking the Ancient Sulfur Cycle. Annual Reviews of Earth Planet Science, 43: 593–622

Finegan P, León Vintró L, Mitchell PI, Boust D, Gouzy A, Kershaw PJ, Lucey JA (2009) Accumulation, solid partitioning and remobilisation of ^{99}Tc in subtidal and intertidal sediments in the Irish Sea. Continental Shelf Research, 29(16): 1995–2010

Francis AJ, Dodge CJ (2008) Bioreduction of Uranium(VI) Complexed with Citric Aciv by *Clostridia* Affects its Structure and Solubility. Environmental Science and Technology, 42: 8277–8282

Froelich PN, Klinkhammer GP, Bender ML, Luedtke NA< Heath GR, Cullen D, Dauphin P (1979) Early oxidation of organic matter in pelagic sediments of the eastern equatorial Atlantic: suboxic diagenesis. Geochimica et Cosmochimica Acta, 43: 1075–1090

Fu B, Conrad R, Blaser M (2018) Potential contribution of acetogenesis to anaerobic degradation in methanogenic rice field soils. Soil Biology and Biochemistry, 119: 1–10

Galamboš M, Kufčáková J, Rosskopková O, Rajec P (2010) Adsorption of cesium and strontium on nitrified bentonites. Journal of Radioanalytical and Nuclear Chemistry, 283: 803–813

Gardes M, Bruns TD (1993) ITS primers with enhances specificity for basidiomycetes – application to the identification of mycorrhizae and rusts. Molecular Ecology, 2(2): 113–118

Ghurye JS, Cepeda-Espinoza V, Pop M (2016) Metagenomic Assembly: Overview, Challenges and Applications. Yale Journals of Biology and Medicine, 89: 353–362

Gray, J., Jones, S.R., Smith, A.D., 1995. Discharges to the environment from the Sellafield Site, 1951-1992. Journal of Radiological Protection, 15: 99–131.

Gu B, Wu W-M, Ginder-Vogel MA, Yan H, Fields MW, Zhou J, Fendorf S, Criddle CS, Jardine PM (2005) Bioreduction of Uranium in a Contaminated Soil Column. Environmental Science and Technology, 39: 4841–4847

Gutell RR, Larsen N, Woese CR (1994) Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiological Reviews, 58(1): 10-26

Halcrow Group (2013) Ravenglass Estuary Complex Sefton Council.

Hamilton EI, Clarke KR (1984) The recent sedimentation history of the Esk estuary, Cumbria, U.K.: The application of radiochronology. Science of the Total Environment, 35: 325–386. doi:10.1016/0048-9697(84)90011-1

Hayatsu M, Tago K, Saito M (2008) Various players in the nitrogen cycle: Diversity and functions of the microorganisms involved in nitrification and denitrification. Soil Science and Plant Nutrition, 54(1): 33–45

Head IM, Saunders JR, Pickup RW (1998) Microbial evolution, diversity, and ecology: A decade of ribosomal RNA analysis of uncultivated microorganisms. Microbial Ecology, 35(1):

Hedderich R, Whitman WB (2006) Physiology and Biochemistry of the Methane-Producing Archaea. in Dworkin M, Falkow S, Rosenberg E, Schleifer K.-H, Stackebrandt E (eds.) The Prokaryotes: Volume 2: Ecophysiology and Biochemistry. New York, NY: Springer New York, pp. 1050-1079

Henderson GS, de Groot FMF, Moulton BJA (2014) X-ray Absorption Near-Edge Structure (XANES) Spectroscopy. Reviews in Mineralogy & Geochemistry, 78: 75–138

Hetherington JA, Harvey BR (1978) Uptake of Radioactivity by Marine Sediments and Implications for Monitoring Metal Pollutants. Marine Pollution Bulletin, 9: 102–106

Holmes DE, Shrestha PM, Walker DJF, Dang Y, Nevin KP, Woodard TL, Lovley DR (2017) Metatranscriptomic Evidence for Direct Interspecies Electron Transfer between Geobacter and Methanothrix Species in Methanogenic Rice Paddy Soils. Applied and Environmental Microbiology, 83(9): e00223-17

Hooker PJ (1991) The geology, hydrogeology and geochemistry of the Needle's Eye natural analogue site (EUR--13434). Commission of the European Communities (CEC)

Iwai S, Weinmaier T, Schmidt BL, Albertson DG, Poloso NJ, Dabbagh K, DeSantis TZ (2016) Piphillin: Improved Prediction of Metagenomic Content by Direct Inference from Human Microbiomes. PLoS One: 1(11): e0166104. doi:10.1371/journal.pone.0166104

Jamet P, Hooker PJ, Schmitt JM, Ledoux E, Escalier des Orres P (1993) Hydrogeochemical modelling of an active system of uranium fixation by organic soils and sediments (Needle’s Eye, Scotland). Mineralium Desposita, 28: 66–76

Johnston CA, Groffman P, Breshears DD, Cardon ZG, Currie W, Emanuel W, Gaudinski J, Jackson RB, Lajtha K, Nadelhoffer K, Nelson D, Post WM,

Retallack G, Wielopolski L (2004) Carbon cycling in soil. Frontiers in Ecology and the Environment, 2(10): 522–528

Jørgensen BB (1982) Ecology of the bacteria of the sulphur cycle with special reference to anoxic-oxic interface environments. Philosophical Transactions of the Royal Society, 298: 543–561

Kemner KM, Kelly SD (2007) Synchrotron-Based Techniques for Monitoring Metal Transformations. In: Hurst CJ, Crawford RL, Garland JL, Lipson DA, Mills AL, Stetzenbach LD (eds) Manual of Environmental Microbiology. Washington DC

Keogh SM, Aldahan A, Possnert G, Finegan P, León Vintró L, Mitchell PI (2007) Trends in the spatial and temporal distribution of 129I and 99Tc in coastal waters surrounding Ireland using Fucus vesiculosus as a bio-indicator. Journal of Environmental Radioactivity, 95: 23–38. doi:10.1016/j.jenvrad.2007.01.009

Kershaw PJ, Swift DJ, Pentreath RJ, Lovett MB (1983) Plutonium redistribution by biological activity in Irish Sea sediments. Nature, 306: 774–775

Kershaw PJ, Swift DJ, Denoon DC (1988) Evidence of recent sedimentation in the eastern Irish Sea. Marine Geology, 85(1): 1–14

Kershaw PJ, Woodhead DS, Malcolm SJ, Allington DJ, Lovett MB (1990) A Sediment History of Sellafield Discharges. Journal of Environmental Radioactivity, 12: 201–241

Kershaw PJ, Pentreath RJ, Woodhead DS, Hunt GJ (1992) A review of radioactivity in the Irish Sea- a report prepared for the Marine Pollution Monitoring Group. Aquatic Environment Monitoring Report.

Kimber RL, Boothman C, Purdie P, Livens FR, Lloyd JR (2012) Biogeochemical behaviour of plutonium during anoxic biostimulation of contaminated sediments. Mineralogical Magazine, 76(3): 567–578

Kimber RL, Corkhill CL, Amos S, Livens FR, Lloyd JR (2015) Geochemical association of Pu and Am in selected host-phases of contaminated soils from the UK and their susceptibility to chemical and microbiological leaching. Journal of Environmental Radioactivity, 142: 96–102

Klotz MG, Bryant DA, Hanson TE (2011) The microbial sulfur cycle. Frontiers in Microbiology, 2: doi: 10.3389/fmicb.2011.00241

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Põldmaa K, Saag L, Saar I, Schüßler A, Scott JA, Senés C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiß M, Larsson K-H (2013) Towards a unified paradigm for sequence-based identification of Fungi. Molecular Ecology, DOI: 10.1111/mec.12481

Korin T (2018) Microbial ecology of anaerobic biodegradation of benzoate: Microbial communities and processes. PhD thesis, Newcastle University.

Lam P, Kuypers MMM (2011) Microbial Nitrogen Cycling Processes in Oxygen Minimum Zones. Annual Reviews of Marine Science, 3: 317–345

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature Biotechnology, 31: 814–821

Langmuir D (1978) Uranium solution-mineral equilibria at low temperatures with applications to sedimentary ore deposits. Geochimica et Cosmochimica Acta, 42(6): 547–569

Lee SY, Baik MH, Cho H-R, Jung EC, Jeong JT, Choi JW, Lee YB, Lee YJ (2013) Abiotic reduction of uranium by mackinawite (FeS) biogenerated under sulfate-reducing conditions. Journal of Radioanalytical and Nuclear Chemistry, 296(3): 1311–1319

Libes SM (1992) An Introduction to Marine Biogeochemistry. John Wiley and Sons, Hoboken, NJ: pp. 734

Loh PS, Reeves AD, Harvey SM, Overnell J, Miller AEJ (2008) The fate of terrestrial organic matter in two Scottish sea lochs. Estuarine, Coastal, and Shelf Science, 76: 566–579

Lloyd JR (2003) Microbial reduction of metals and radionuclides. FEMS Microbiology Reviews, 27: 411–425

Lloyd JR, Beveridge TJ, Morris K, Polya DA, Vaughan DJ (2007) Techniques for Studying Microbial Transformations of Metals and Radionuclides. In: Hurst CJ, Crawford RL, Garland JL, Lipson DA, Mills AL, Stetzenbach LD (eds) Manual of Environmental Microbiology. Washington DC

Lovley DR, Phillips EJP, Gorby YA, Landa ER (1991) Microbial reduction of uranium. Letters to Nature, 350: 413–416

Lovley DR, Roden EE, Phillips EJP, Woodward JC (1993) Enzymatic iron and uranium reduction by sulfate-reducing bacteria. Marine Geology, 113: 41–53

Lovley DR (2011) Live wires: direct extracellular electron exchange for bioenergy and the bioremediation of energy-related contamination. Energy and Environmental Science, 4: 4896–4906

MacKenzie AB, Scott RD, Houston CM, Hooker PJ (1991) Natural decay series radionuclide studies at the Needle's Eye natural analogue site (EUR—13126). Commission of the European Communities (CEC)

MacKenzie AB, Cook GT, McDonald P, Jones S (1998) The Influence of Mixing Timescales and Re-dissolution Processes on the Distribution of Radionuclides

in Northeast Irish Sea Sediments. Journal of Environmental Radioactivity, 39: 35–53

Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA (2018) Brock biology of microorganisms (Fifteenth edition.). Boston: Pearson.

Malcolm SJ, Kershaw PJ, Cromar NJ, Botham L (1990) Iron and managense geochemitrsy and the distribution of ^{239,240}Pu and ^{241}Am in the sediments of the North East Irish Sea. The Science of the Total Environment, 95: 69–87

Malik S, Beer M, Megharaj M, Naidu R (2008) The use of molecular techniques to characterize the microbial communities in contaminated soil and water. Environment International, 34(2): 265–276

Markley JL, Brüschweiler R, Edison AS, Eghbalnia HR, Powers R, Raferty D, Wishart DS (2017) The future of NMR-based metabolomics. Current Opinion in Biotechnology, 43: 34–40

Marsden OJ, Abrahamsen L, Bryan ND, Philip Day J, Keith Fifield L, Gent C, Goodall PS, Morris K, Livens FR (2006) Transport and accumulation of actinide elements in the near-shore environment: field and modelling studies. Sedimentology, 53: 237–248. doi:10.1111/j.1365-3091.2005.00761.x

McBeth JM, Lear G, Lloyd JR, Livens FR, Morris K, Burke IT (2007) Technetium Reduction and Reoxidation in Aquifer Sediments. Geomicrobiology Journal, 24: 189–197

McCann CM (2012) The use of natural Mn oxide-containing wastes as a contaminated land remediation strategy and their effects on soil microbial functioning. PhD thesis, Newcastle University, viewed 1 June

McCartney M, Kershaw PJ, Woodhead DS, Denoon DC (1994) Artificial radionuclides in the surface sediments of the Irish Sea, 1968–1988. Science of the Total Environment, 141: 103–138. doi:10.1016/0048-9697(94)90022-1

McGrath KC, Thomas-Hall SR, Cheng CT, Leo L, Alexa A, Schmidt S, Schenk PM (2008) Isolation and analysis of mRNA from environmental microbial communities. Journal of Microbiological Methods, 75(2): 172–176

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens A, Wilke A, Wilkening J, Edwards RA (2008) The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. Bioinformatics, 9(386): https://doi.org/10.1186/1471-2105-9-386

Miller JM, Taylor K (1966) Uranium mineralization near Dalbeattie, Kirkcudbrightshire. Bulletin of the Geological Survey of Great Britain, 25: 1–18

Milodowski AE, West JM, Pearce JM, Hyslop EK, Basham IR, Hooker PJ (1990) Uranium-mineralized microogranisms associated with uraniferous hydrocarbons in southwest Scotland. Nature, 347: 465–467

Morris K, Butterworth JC, Livens FR (2000) Evidence for the Remobilization of Sellafield Waste Radionuclides in an Intertidal Salt Marsh, West Cumbria, U.K. Estuarine, Coastal and Shelf Science, 51: 613–625. doi:10.1006/ecss.2000.0705

Morse JW, Choppin GR (1991) The Chemistry of Transuranic Elements in Natural Waters. Reviews in Aquatic Science, 4: 1–22

Muir G, Cook GT, Tripney BG, MacKenzie AB, Stewart H, Tierney KM (2015) Temportal Trend in the Transfer of Sellafield-Derived ^{14}C into Different Size Fractions of the Carbonate Component of the NE Irish Sea Sediment. Radiocarbon, 57(3): 347–354

Muyzer G (1999) Genetic fingerprinting of microbial communities: present status and future perspective, in: Bell R, Brylinsky M, Johnson-Green P (eds) Proceedings of the 8th International Symposium on Microbial Ecology. Atlantic Canada Society for Microbial Ecology: Halifax, Nova Scotia., pp. 1-10

Neidhardt FC, Ingraham J, Low KB, Magasanik B, Schaechter M, Umbarger HE (1996) Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology, Volumes 1 & 2. Microbiology, Washington DC.

Newsome L, Morris K, Lloyd JR (2014) The biogeochemistry and bioremediation of uranium and other priority radionuclides. Chemical Geology, 363: 164–184

Olsen GJ, Lane DJ, Giovannoni SJ, Pace NR, Stahl DA (1986) Microbial ecology and evolution: a ribosomal RNA approach. Annual Review of Microbiology, 40 337-365

Op de Beeck M, Lievens B, Busschaert P, Declerck S, Vangronsveld J, Colpaert JV (2014) Comparison and Validation of Some ITS Primer Paires Useful for Fungal Metabarcoding Studies. PLoS ONE, 9(6): e97629. https://doi.org/10.1371/journal.pone.0097629

Payne RB, Gentry DM, Rapp-Giles BJ, Casalot L, Wall JD (2002) Uranium Reduction by *Desulfovibrio desulfuricans* Strain G20 and a Cytochrome c_3 Mutant. Applied and Environmental Microbiology, 68(6): 3129–3132

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41(D1): 590–596

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N (2017) Shotgun metagenomics, from sampling to analysis. Nature Biotechnology, 35(9): 833–845

Ragsdale SW, Pierce E (2008) Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation. Biochimica et Biophysica Acta – Proteins and Proteomics, 1784(12): 1873–1898

Rastogi G, Osman S, Vaishampayan PA, Andersen GL, Stetler LD, Sani RK (2010) Microbial Diversity in Uranium Mining-Impacted Soils as Revealed by High-Density 16S Microarray and Clone Linrary. Microbial Ecoloy, 59: 94–108

Renshaw JC, Butchin LJC, Livens FR, May I, Charnock JM, Lloyd JR (2005) Bioreduction of Uranium: Environmental Implications of a Pentavalent Intermediate. Environmental Science and Technology, 39: 5657–5660

Renshaw JC, Law N, Geissler A, Livens FR, Lloyd JR (2009) Impact of the Fe(III)-reducing bacteria *Geobacter sulfurreducens* and *Shewanella oneidensis* on the speciation of plutonium. Biogeochemistry, 94: 191–196

Ridgway IM, Price NB (1987) Geochemical associations and post-depositional mobility of heavy metals in coastal sediments: Loch Etive, Scotland. Marine Chemistry, 21: 229–248

Rollinson HR (1993) Using Geochemical Data: Evaluation, Presentaion, Interpretation. Routledge, New York

Rusin PA, Quintana L, Brainard JR, Strietelmeier BA, Talt CD, Ekberg SC, Palmer PD, Newton TW, Clark DL (1994) Solubilization of Plutonium Hydrous Oxide by Iron-Reducing Bacteria. Environmental Science and Technology, 28: 1686–1690

Saenen E, Horemans N, Vanhoudt N, Vandenhove H, Biermans G, van Hees M, Wannijn J, Vangronsveld J, Cuypers A (2013) Effects of pH on uranium uptake and oxidative stress responses induced in *Arabidopsis thaliana*. Environmental Toxicology and Chemistry, 32: 2125-2133

Schloss PD, Wescott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, van Horn DJ, Weber CF (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology, 75(23): 7537–7541

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proceedings of the National Academy of Sciences, 109(16): 6241–6246

Scholz MB, Lo C-C, Chain PSG (2011) Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. Current Opinion in Biotechnology, 23: 9–15

Scott RD, MacKenzie AB, Ben-Shaban YA, Hooker PJ, Houston CM (1991) Uranium Transport and Retardation at the Needle's Eye Natural Analogue Site, South West Scotland. Radiochimica Acta, 52/53: 357–365

Sedlar K, Kupkova K, Provaznik I. (2017) Bioinformatics strategies for taxonomy independent binning and visualization of sequences in shotgun metagenomics. Computational and Structural Biotechnology Journal, 15: 48–55

Selinger DW, Saxena RM, Cheung KJ, Church GM, Rosenow C (2003) Global NRA Half-Life Analysis in Escherichia coli Reveals Positional Patterns of Transcript Degradation. Genome Research, 1(3): 216–223

Shah N, Tang H, Doak TG, Ye Y (2011) Comparing bacterial communities inferred from 16S rRNA gene sequencing and shotgun metagenomics. Biocomputing 2011: 165–176

Sharma S, Mehta R, Gupta R, Schloter M (2012) Improved protocol for the extraction of bacterial mRNA from soils. Journal of Microbiological Methods, 91(1): 62–64

Singh G, Sevinç Şengör S, Bhalla A, Kumar S, De J, Stewart B, Spycher N, Ginn TM, Peyton BM, Sani RK (2014) Reoxidation of Biogenic Reduced Uranium: A Challenge Toward Bioremediation. Critical Reviews in Environmental Science and Technology, 44: 391–415

Spiro S, Guest JR (1990) FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiology Reviews, 6(4): 399–428

Staley JT, Konopka A (1985) Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annual Review of Microbiology, 39: 321-346

Sutcliffe B, Charlton AA, Harford AJ, Hose GC, Greenfield P, Elbourne LDH, Oytam Y, Stephenson S, Midgley DJ, Paulsen IT (2017) Effects of uranium concentration on microbial community structure and functional potential. Environmental Microbiology, 19(8): 3323–3341

Suzuki Y, Suko T (2006) Geomicrobiological factors that control uranium mobility in the environment: Update on recent advances in the bioremediation of uranium-contaminated sites. Journal of Mineralogical and Petrological Sciences, 101: 299–307

Tang J (2011) Microbial Metabolomics. Current Genomics, 12(6): 391–403

Thauer RK, Jungermann K, Decker K (1977) Energy Conservation in Chemotrophic Anaerobic Bacteria. Bacteriological Reviews, 41: 100–180, doi:10.1073/pnas.0803850105

Thomas R (1993) Practical Guide to ICP-MS. 1st ed. Marcel Dekker, New York

Thomas T, Gilbert J, Meyer F (2012) Metagenomics – a guide from sampling to data analysis. Microbial Informatics and Experimentation, 2:3

Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-Coverage ITS Primers for the DNA-Based Identification of Ascomycetes and Basidiomycetes in Environmental Samples. PLoS ONE, 7(7): e40863, doi.org/10.1371/journal.pone.0040863

Tremaroli V, Workentine ML, Weljie AM, Vogel HJ, Ceri H, Viti C, Tatti E, Zhang P, Hynes AP, Turner RJ, Zannoni D (2009) Metabolomic Investigation of the Bacterial Response to a Metal Challenge. Applied and Environmental Microbiology, 75(3): 719–728

Unden G, Becker S, Bongaerts J, Holighaus G, Schirawski J, Six S (1995) O_2-Sensing and O_2-dependent gene regulation in facultatively anaerobic bacteria. Archives of Microbiology, 164(2): 81–90

van de Peer Y, Chapelle S, de Wachter R (1996) A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucleic Acids Research, 24(17): 3381-3391

White TJ, Bruns T, Lee SJ, Taylor JL (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications, 18(1):315-22

Williams KH, Bargar JR, Lloyd JR, Lovley DR (2013) Bioremediation of the uranium-contaminated groundwater: a systems approach to subsurface biogeochemistry. Current Opinion in Biotechnology, 24: 489–497

Williams MA, Taylor EB, Mula HP (2010) Metaproteomic characterization of a soil microbial community following carbon amendment. Soil Biology and Biochemistry, 42(7): 1148–1156

Woese CR (1987) Bacterial evolution. Microbiological Reviews, 51(2): 221-271

Wood BJB, Tett PB, Edwards A (1973) An Introduction to the Phytoplankton, Primary Production and Relevant Hydrography of Loch Etive. Journal of Ecoloy, 61(2): 569–585

Xia Y, Wang Y, Fang HHP, Jin T, Zhong H, Zhang T (2014) Thermophilic microbial cellulose decomposition and methanogenesis pathways recharacterized by Metatranscriptomic and metagenomic analysis. Scientific Reports, 4: DOI: 10.1038/srep06708

Yan X, Luo X (2015) Radionuclides distribution, properties, and microbial diversity of soils in uranium mill tailings from southeastern China. Journals of Environmental Radioactivity, 139: 85–90

Yan Y, Luo X, Xhao M (2016) Metagenomic analysis of microbial community in uranium-contaminated soils. Applied Microbiology and Biotechnology, 100: 299–310

Zhang YJ, Bryan ND, Livens FR, Jones MN (1997) Selectivity in the complexation of actinides by humic substances. Environmental Pollution, 96(3): 361–367

Chapter 2. Methods and Materials

2.1 Sample Collection

2.1.1 Needle's Eye

Two soil depth cores were collected in December 2014 and December 2015 from two separate bogs located approximately 200 metres apart, from the Needle's Eye site (Dumfries, UK) (54°53'15" N, 3°41'35" W) (Figure 1.1). One core was taken as a representative of the contaminated bog, and the second core was taken from nearby similar location as a putative control, under the assumption it shared similar geochemical characteristics but had potentially lower uranium (U) levels. This assumption was later confirmed by ICP-MS (see Chapter 3). The contaminated core, collected in December 2014, was taken using an acrylic tube that was capped at the top. The tube was pushed into the bog soil, raised, and capped at the bottom as soon as possible. The decision to sample the contaminated site this way was made with the intention of retaining *in situ* depth dependent redox conditions and to maintain porewater interface profiles, both to aid geochemical analysis. The contaminated core was immediately returned to the University of Manchester and was sampled the following day in an O_2-free, N_2-filled Atmosbag (Sigma, UK) on a coring table. The core was sampled at 1 cm resolution to a depth of 41 cm, with soil samples taken simultaneously for separate geochemical and microbial analysis and frozen at -20°C. Due to the method used to originally extract the core from the site, it was not viable to extract porewater from each depth for analysis—therefore, an alternative core was collected at later date by Fuller *et al.* (in prep.) to identify pore water variations in geochemistry. The samples intended for microbial analysis were frozen at -20°C and returned to Newcastle University the following day and stored at -20°C until they were processed for DNA extraction within one week.

In December 2015, the control core was taken using an auger corer in 10 cm sections, with each section wrapped in cellophane, placed in a cool-box, and returned to Newcastle University, and stored overnight at -20°C. The following day, the 10 cm sections were sampled at 1 cm resolution to a depth of 41 cm, whereupon the DNA was extracted immediately.

2.1.2 Ravenglass and Irish Sea Mud Patch

Two depth cores were collected in September 2014 from a saltmarsh in the intertidal zone of the River Esk estuary, near the town of Ravenglass (Cumbria, UK) (54°20'24" N, 3°24'06" W) (Figure 1.3). A core tube was pushed into the sediment, capped at the top, extracted, and capped at the bottom before complete removal, to maintain porewater/sediment interface (*i.e.*, porewater collected from the same depth as sediment) profiles and anaerobic conditions. The Ravenglass core was returned to the University of Manchester and sampled in an O_2-free, N_2-filled Atmosbag on a coring table. The core was sectioned at 1 cm resolution between 1 – 10 cm, and at 2 cm resolution thereafter, to a depth of 28 cm. Sediment was collected for both geochemical and microbial analysis, and frozen upon sectioning.

A depth core was also collected in July 2014 from a belt of fine-grained sediment in the north east Irish Sea mud patch (54°26'80" N, 03°42'89" W) (Figure 1.3). Acrylic core tubes were inserted into the sediment using a hydraulic megacorer (Law *et al.*, 2009). Cores were capped at the bottom upon retrieval. The cores were sampled within one hour of retrieval and to a depth of 36 cm. Frozen samples from all cores were returned to the University of Manchester and stored at -20°C, before being transported to Newcastle University at a later date and stored again at -20°C. DNA was extracted from all samples in January 2015.

2.1.3 Loch Etive

One depth core was collected each from two separate parts of Loch Etive (56°26'90" N, 5°14'77" W) near the town of Oban (Scotland, UK) in June 2017 (Figure 1.4). One core was collected from a shallow basin at Airds Bay of Loch Etive, from a water depth of 71 metres. The other core was collected from a deep basin of Loch Etive, from a water depth of 147 metres. Acrylic tubes were inserted into the basin sediment using a hydraulic megacorer as per the Irish Sea mud patch and capped upon retrieval. Cores were taken to the Scottish Association for Marine Science (SAMS) laboratory near Oban and stored at 4°C overnight. The following day, the cores were sampled in an O_2-free, N_2-filled Atmosbag on a coring table, at 0.5 cm resolution between 1 – 2 cm, 1 cm resolution between 1 – 12 cm and 1 – 10 cm, 2 cm resolution between 14 – 20 cm and 12 – 22 cm, 5 cm resolution between 25 – 35 cm and 27 – 47 cm for the shallow basin and deep basin respectively. The

shallow basin had a final sample comprised of sediment from 35 – 42 cm. Sediment from each of the depth increments was taken for both geochemical and microbial analysis and returned to the University of Manchester and Newcastle University respectively, in a cool-box and upon samples were stored at -20°C. DNA was extracted from the cores within a week of retrieval.

2.2 DNA Extraction and Quantification

DNA was extracted from each depth sample from all the cores described above by taking ~ 500 mg of soil or sediment and using the FastDNA Spin Kit for Soil (MP Biomedicals, USA) as per manufacturer's instructions. A summary of the extraction process is as follows: Firstly, ~ 500 mg of soil was placed into a weighed Lysing Matrix E tube (containing a mix of lysing beads) and reweighed. Then 978 μL of Sodium Phosphate Buffer and 122 μL of MT Buffer was added. Samples were then homogenised (speed 6.0 on the FastPrep-24) and centrifuged at 14,000 x *g* for 10 minutes. Supernatant containing the lysed DNA was placed into another 2 mL tube, and 250 μL of Protein Precipitation Solution was added and mixed, followed by centrifugation at 14,000 x *g* for five minutes. Supernatant (1 mL) and 1 mL proprietary binding matrix was then placed in a 2 mL tube, mixed, and left to settle after which 500 μL of supernatant was removed. The remaining supernatant was re-suspended, and placed in 2 mL tubes with spin filters, which were used to trap the bound DNA and remove the supernatant. Bound DNA on the filter was cleaned with SEWS-M, a proprietary wash solution containing ethanol that removes impurities from matrix-bound DNA, and centrifuged for 2 minutes at 14,000 x *g*. DNA was then eluted from the filter by adding 50 μL of ultra-sterile water and centrifuging for 1 minute at 14,000 x *g*. All samples from an individual core were extracted at the same time, and a procedural blank extraction (comprised of 500 μL of molecular grade water) was performed additionally for each core. Extracted template DNA was first diluted (1:10) with molecular grade water to increase template volume and to decrease the levels of any PCR-inhibiting substances that may have been co-extracted. Diluted template DNA was then quantified using a Qubit 3.0 fluorometer using the dsDNA HS Assay Kit according to the manufacturer's instructions (Life Technologies, USA). Extracted DNA was stored at 4°C when in use and at -20°C for longer-term storage.

2.3 'End Point' PCR of Functional, 16S, and ITS Genes

A series of primer pairs were chosen to target a range of both taxonomic marker and functional gene fragments representative of microbial catabolic processes. The primer pairs and respective PCR protocols used, and associated references are shown in Table 2.1. Two primer sets targeted the 16S rRNA gene, with one intended as a universal primer pair that covered all bacteria and archaea, and a second that targeted the *Geobacter* spp. exclusively. These 'end point' PCR amplifications were used to give a qualitative indication of the target gene abundances within the depth cores, and to direct subsequent sample selection for sequencing. The taxonomic marker genes targeted were the prokaryotic 16S rRNA gene (V4-V5 region) and the fungal ITS2 rRNA gene. Primers targeting specifically the bacterial genus *Geobacter* spp. were used to interpret abundance of this particular genus within depth samples, as it is heavily implicated in heavy metal and actinide bioreduction (Lovley *et al.*, 1993). Targeted functional genes were chosen based on their catabolic function to provide a range of functional processes indicative of a variety of redox conditions, from highly oxidising to highly reducing. Genes that encode for enzymes that are critical and specific to certain microbial respiratory pathways were targeted with primers, and the intensity of the PCR amplicon band was interpreted as an analogy for gene abundance within a given sample. The genes targeted were: ammonia monooxygenase (*amoA*), separately for both bacteria and archaea; nitrite reductase (*nirK* and *nirS*); dissimilatory sulphite reductase (*dsrA*), and; methyl CoM reductase (*mcrA*). These four genes encapsulated a broad range of microbial respiratory processes from a redox perspective *i.e.*, from highly oxidising to highly reducing. There are no known genes that are specific to microbial iron reduction, a process which is essential to the biogeochemical cycling of iron. Furthermore, since iron reduction is closest to uranium reduction in terms of E_h this microbially mediated process would have been a particularly relevant functional target. In lieu of this, it was decided that targeting genes from denitrification and dissimilatory sulfate reduction would suffice to 'bookend' these metal reduction processes, *i.e.*, patterns of denitrification abundance shifting to sulfate reduction abundance infers the occurrence of metal reduction in-between. Procedural positives were obtained for each target gene and are included in Table 2.2.

All ‘end point’ PCR amplifications were performed on a Techne 512 thermocycler. Each PCR reaction contained 10 μL, comprised of: 9 μL of MegaMix Blue (Microzone, UK), itself containing *taq* polymerase, dNTPs, $MgCl^{2-}$, buffer, and agarose loading dye; 0.25 μL of each primer (100 pM concentration), and; 0.5 μL of 1:10 diluted template DNA. No concession was made to the differing quantities of template DNA within each depth sample. PCRs were run as per the protocols for each primer described in Table 2.1. Procedural blanks were run for each reaction batch, using 0.5 μL of molecular grade water instead of template DNA.

After PCR, amplicons were checked for product and fragment size using agarose gel electrophoresis. 5 μL of amplicon products was loaded onto a 1% w/v agarose gel, comprised of 1 g of agarose (UltraPure, Invitrogen, USA), 99 mL of 1 x TAE (tris-acetate-EDTA) buffer (40 mM tris, 20 mM acetic acid, 1 mM EDTA, pH 8.3), and 1.6 μL of ethidium bromide (Bio-Rad, USA) for staining. Since PCR products already contained loading dye as part of the MegaMix Blue reagent, no additional dye was added. Marker lanes were loaded with 2.5 μL of HyperLadder 50bp (BioLine, UK), which contained DNA fragments of varying sizes, ranging from 50 bp to 2.1 kbp, to determine PCR amplicon fragment size. Electrophoresis was performed for 45 minutes at 100 V. Gels were visualised using a UV MultiDoc-It™ Imaging System (UVP, USA). Gel images were inspected and analysed using the BioNumerics (Applied Maths, USA) software, where each PCR amplicon fragment band in a gel was given a grey-scale 8-bit value (0 – 256-pixel count) based on the densitometric curve analysis of individual bands. The purpose of this ‘end point’ PCR method was to identify variation in the amount of amplicon for a given gene target as a function of depth in respective soil and sediment cores. Critically, a higher pixel value, *i.e.*, band intensity as derived from gel analysis in BioNumerics, was presumed to indicate a larger gene abundance in the original template DNA, despite quantities of template DNA being shown to be different between samples as a systematic function of depth.

2.4 DGGE Preparation and Analysis

Denaturing gradient gel electrophoresis was used to determine differences in microbial communities between samples using 16S and ITS rRNA gene primers. PCRs were performed as above but utilised primers with a GC-clamp (5` –

Table 2.1. List of functional and marker genes targeted by 'end point' PCR, description of the gene function, name of the primers, sequences of the primers from 5` to 3`, the annealing site where available, the target fragment length, the PCR cycling conditions, and the source of the primers and also PCR cycle conditions where appropriate.

Gene	Primer Name	Sequence (5' -> 3')	Annealing Site	Fragment Length	Cycle Details	Reference
AOB *amoA*	amoA-1f	GGGGHTTYTACTGGTGGT	332-349[1]	530	94C for 10 min 94C for 1 m x 42 57C for 1 m x 42 72C for 1 m x 42 72C for 15 m	Stephen *et al.* (1999) McCann (2012)
	amoA-2r	CCCCTCKGSAAAGCCTTCTTC	802-822[1]			
AOA *amoA*	amoA-1f	STAATGGTCTGGCTTAGACG	N/A[7]	635	95C for 10 m 94C for 45 s x 30 55C for 1 m x 30 72C for 1 m x 30 72C for 15 m	Francis *et al.* (2005) McCann (2012)
	amoA-2f	GCGGCCATCCATCTGTATGT	N/A[7]			
nirK	FlaCu	ATCATGGTSCTGCCGCG	568-584[2]	512	94C for 2 m 94C for 30 s x 30 57C for 1 m x 30 72C for 1 m x 30 72C for 10 m	Throbäck *et al.* (2004) McCann (2012)
	R3cu	GCCTCGATCAGRTTGTGGTT	1021-1040[2]			
nirS	Cd3af	GTSAACGTSAAGGARACSGG	916-935[3]	465	94C for 2 m 94C for 30 s x 35 57C for 1 m x 35 72C for 1 m x 35 72C for 10 m	Throbäck *et al.* (2004) McCann (2012)
	R3cd	GASTTCGGRTGSGTCTTGA	1322-1341[3]			
dsrA	DSR1F	ACSCAYTGGAARCACG	N/A[7]	c. 1900	94C for 1 m 94C for 40 s x 30 54C for 45 s x 30 72C for 2 m x 30 72C for 10 m	Klein *et al.* (2001) Guan *et al.* (2013)
	DSR4R	GTGTARCAGTTDCCRCA	N/A[7]			
mcrA	ML-f/mcrF	GGTGGTGTMGGATTCACACA--RTAYGCWACAGC	N/A[7]	460	95C for 5 m 95C for 1 m x 35 49C for 1 m x 35 72C for 3 m x 35 72C for 10 m	Luton *et al.* (2002) Morris *et al.* (2014)
	ML-r/mcrR	TTCATTGCRTAGTTWGGRTA--GTT	N/A[7]			
ITS2	ITS3_KYO2	GATGAAGAACGYAGYRAA	2024-2043[4]	435	95C for 10 m 95C for 1 m x 30 55C for 1 m x 30 72C for 1 m x 30 72C for 10 m	Gardes and Bruns (1993) Toju *et al.* (2012)
	ITS4	TCCTCCGCTTATTGATATGC	2390-2409[4]			
16S rRNA Universal	515F	GTGNCAGCMGCCGCGGTAA	515-534[5]	450	95C for 10 m 95C for 1 m x 30 55C for 1 m x 30 72C for 1 m x 30 72C for 10 m	Muyzer *et al.* (1993)
	926R	CCGYCAATTYMTTTRAGTTT	926-946[5]			
16S rRNA *Geobacter* specific	Geo564F	AAGCGTTGTTCGGAWTTAT	564-583[6]	246	94C for 4 m 94C for 30 s x 35 65-55C for 30 s x 35 72C for 30 s x 35 72C for 3 m TD of 0.5C for 20 steps	Cummings *et al.* (2003)
	Geo840R	GGCACTGCAGGGGTCAATA	840-859[6]			

[1] Positions in the *amoA* gene sequence of *Nitrosomonas europaeae* (McTavish *et al.*, 1993)
[2] Positions in the *nirK* gene sequence of *Alcaligenes faecalis* S-6 (Throbäck *et al.*, 2004)
[3] Positions in the *nirS* gene sequence of *Psuedomonas stutzeri ZoBell* ATCC 14405 (Throbäck *et al.*, 2004)
[4] Positions in the ITS rRNA gene sequence of *Serpula himantiodes* (AM946630) (Toju *et al.*, 2012)
[5] Positions in the 16S rRNA gene sequence of *Escherichia coli* (Brosius *et al.*, 1978)
[6] Positions in the 16S rRNA gene sequence of *E. coli* (Brosius *et al.*, 1978)
[7] Positions not described in source

Table 2.2. List of PCR functional gene procedural positive DNA sources.

Gene	Procedural Positive DNA	Source
amoA – AOB	Sequenced Soil DNA from St Anthony's Lead Works, Newcastle upon Tyne.	McCann (2012)
amoA – AOA		
nirK		
nirS		
dsrA	*Desulfobulbus*-like Clone	Jarvis *et al.* (2014)
mcrA	*mcrA* plasmid	*E. coli* clone
ITS2	*Trichoderma Reesei*	DSMZ – No. 768
16S rRNA	*E. coli*	DSMZ – No. 10198
Geobacter spp.	N/A	N/A

CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCCC – 3`; Muyzer *et al.*, 1993) attached at the 5` end of the forward primer, to increase the stability of this end of the PCR amplicon within the denaturing gradient. All DGGEs were performed on a D-CODE System (Bio-Rad, USA). DGGE gels used a 30:70 gradient, *i.e.*, the gel is denser at the bottom, to retard migration of shorter fragments, using either a 10% polyacrylamide concentration for bacterial 16S rRNA amplicons or 6% concentration for fungal ITS2 amplicons—this variation was to maximise migration of the amplicons through the gel based on fragment size and charge. Gel solutions were prepared separately for the high and low gradients, by adding urea (Sigma, UK) (10.08 g and 5.04 g for high and low concentrations for the gradient respectively), polyacrylamide (acrylamide-N, N- methylenebisacrylamide ratio, 37.5:1; Sigma, UK) depending on gel concentration (10 mL for 10%, 6 mL for 6%), deionised formamide (ThermoFisher Scientific, USA) (9.6 mL for high, 4.8 mL for low concentration for gradient), 0.8 mL of 50 x TAE buffer, with both high and low concentrations made up to a final volume of 40 mL by adding molecular grade water. Finally, 160 μL of 0.1 wt% APS (ammonium persulfate) and 16 μL of TEMED (Tetramethylethylenediamine) was added to both high and low concentrations to begin the gel-setting. Gels were poured into 16 x 16 cm plates (0.75 mm thick) by a peristaltic pump (Model 520, Watson Marlow, UK) and a gradient former (Model 485, Bio-Rad, UK). 20-lane combs were added to each gel before they set. 11 μL of each PCR amplicon product, though they already contained loading dye as part of the MegaMix Blue reagent, had an additional 11 μL of 6 x loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol in filter sterile water; Sigma, UK) added. 5 μL of HyperLadder 50 bp (Bioline, UK) was loaded directly into three or four of the lane wells, to be used as gel markers, to allow for gel normalisation across lanes of the gels. Gels were run at either 200 V at 60°C for 4

hours or 50 V at 60°C for 12 hours, in ~ 7 L of 1 x TAE buffer. Following electrophoresis, gels were stained with 50 mL of 1 x TAE and 9 μL of SYBR Gold (Sigma, UK) and placed in the dark for 30 minutes. Additional 1 x TAE was added to ensure an even spread of the stain, by gently shaking the gels for an additional 30 minutes. Gels were visualised using a UV MultiDoc-It™ Imaging System (UVP, USA). Gel images were inspected and analysed using the BioNumerics (Applied Maths, USA) software.

Images of DGGEs created in BioNumerics, were corrected for any irregularities introduced during the pouring and setting of the gels by normalising gels based on the marker lane fragments. This also allowed for normalisation between different gels of the same percentage, gradient, and PCR amplicon target type. Individual bands within gel lanes were interpreted to represent individual biological representative units, and bands in the same location relative to the marker lane in different lanes were treated to be the same biological unit, also referred to as operational taxonomic units (OTUs). Bands were identified in BioNumerics both by using the 'band detect' feature (set to 1% threshold) and by additional manual identification of bands. Bands between lanes that were considered to be the same biological unit were identified using the 'band match' function in BioNumerics. Band intensity, as determined by the densitometric curve of the band on an 8-bit scale (0–256) was used to quantify band abundance within a lane. Band quantifications of the entire gel was exported from BioNumerics for subsequent statistical analysis.

Multivariate statistical analysis of DGGE data was performed in PRIMER V6 (Clarke and Gorley, 2006). Band abundance data was first standardised by total abundance of all bands within a lane (sample) and square root transformed to reduce the skew induced by large discrepancies of very highly abundant and very low abundant bands. Dissimilatory matrices were then produced using either the Bray-Curtis equation (taking abundance of bands into account) or Jaccard (presence/absence). These matrices were then vectorised, and 2D non-Metric Multi-Dimensional Scaling (nMDS) ordination plots produced. Cluster analysis was also performed on the matrices to identify the percentage similarity grouping between samples. Cluster analysis results were then presented on the 2D nMDS plots to illustrate percentage similarity of samples, at 20, 40, or 60% similarity. The nMDS plots with cluster analyses were used with an aim to discuss overall trends in beta diversity of microbial communities between samples. However, subsequent to

running DGGEs where PCR targeted the 16S rRNA gene, it became clear that the microbial communities were far too diverse to accurately make inferences about community structure as a function of depth *via* the DGGE method, *i.e.*, there were far too many bands to be confident in discerning what bands represented separate biological units, and which bands were the same biological unit between samples. Fungal communities were far less diverse and a more confident description of the fungal communities within the cores could be deduced from this approach.

2.5 DNA Sequencing using next generation sequencing platforms

2.5.1 16S rRNA Gene Sequencing

2.5.1.1 *Ion Torrent Sequencing*

Samples were prepared for sequencing on the Ion Torrent platform (ThermoFisher Scientific, USA) by undergoing PCR using the 515F and 926R primers targeting the V4-V5 region of the 16S rRNA gene (*Escherichia coli* numbering). Forward primers contained Torrent Adaptor A, 12 nt Golay barcodes (Hamady *et al.*, 2008), and a 4-nucleotide key sequence (TCAG) to mark the beginning of the targeted sequence. The final amplicon fragment length was 481 nt. PCRs were 25 μL reactions comprised of: 22 μL of MegaMix Blue; 1 μL of forward and 1 μL of reverse primer (1pp pM), and; 1 μL of 1:10 diluted template DNA. Initial denaturation was 95°C for 5 minutes, followed by 30 cycles of 95°C for 1 minute, 55°C for 1 minute, 72°C for 1 minute, and a final elongation step of 72°C for 10 minutes. Completed PCR reactions were held at 4°C. Each depth sample was run in triplicate to increase amplicon product volume, and each barcode was run with a procedural positive (genomic DNA from *Escherichia coli*) and blank (molecular grade water). PCR amplicons were product checked *via* agarose gel electrophoresis as above. Triplicate PCR reactions were then pooled, and procedural positives and blanks discarded. Pooled amplicons were cleaned using Agencourt AMPure XP (Beckman Coulter, USA). Firstly, a specific volume of pooled triplicate PCR product was taken, usually 50 μL, and place in a 1.5 mL tube, followed by 1.1 x volume (55 μL) of AMPure magnetic beads (1.1 x was determined by the optimal volume of AMPure for the desired PCR fragment length), where the AMPure contains magnetic beads that are used to bind to the DNA; the mixed sample was then incubated at room temperature for 5 minutes on a DynaMag-2 magnetic rack (Life Technologies

Inc, USA); DNA bound by magnetic beads is attracted to the magnetic rack and so sticks to the side of the tube, while the remaining supernatant containing excess PCR reagents and non-target DNA was removed; the DNA beads were cleaned with 80 μL of 70% ethanol, which was removed, and the step repeated. Beads containing DNA were air-dried at room temperature for 5 minutes, to allow excess ethanol to evaporate; DNA was then separated from the magnetic beads by adding 20 μL of molecular grade water on a non-magnetic rack. This entire clean-up process was performed twice to ensure all unwanted PCR reagents and untargeted DNA was removed. Cleaned PCR amplicons were quantified using a Qubit 3.0 fluorometer (as above) and a total dilution factor was calculated for each sample, which were then diluted down to 1000 pM by adding a calculated volume of molecular grade water. Lastly, an identical volume (10–20μL depending on the number of individual samples comprising the final pooled library) was taken from each sample to create a pooled equimolar multiplexed amplicon library which was submitted for sequencing in-house at Newcastle University. Prior to sequencing, pooled libraries were quantified using a Qubit 3.0 fluorometer, and target amplicon fragment length was selected by using the Pippin Prep DNA size selector (Sage Science, USA) gel electrophoresis kit. Template preparation and enrichment of DNA was performed on an Ion OT 2 400 Kit (ThermoFisher Scientific, USA) as per manufacturer's instructions. Sequencing was performed internally using an Ion Torrent Personal Genome Machine on a 316 chip. A multiplexed FASTQ file containing single-end reads was produced for subsequent pipeline analysis.

2.5.1.2 *Illumina MiSeq Sequencing*

In addition to Ion Torrent sequencing, diluted template DNA extractions were submitted for sequencing on the Illumina MiSeq platform (Illumina, USA) at NUOMICs group (Northumbria University). Here samples first underwent PCR using the 515F and 806R primers, which target the V4 region of the 16S rRNA gene (*Escherichia coli* numbering). Primers contained indexes specific to the Illumina platform, and Golay barcodes for subsequent individual sample identification and demultiplexing from pooled amplicon libraries. PCR reactions contained 17 μL of Accuprime Pfx Supermix (ThermoFisher Scientific, USA), 2 μL of each primer (100 μM), and 1 μL of template DNA. Procedural controls (1 μL of molecular grade water) and positive (1 μL of Mock Community) were also run. PCR plates were placed in a thermocycler and underwent the following PCR conditions: Initial denaturation at

95°C for 2 minutes, followed by 30 cycles of; 95°C for 20 seconds, 55°C for 15 seconds, and 72°C for 5 minutes, and; a final elongation step at 72°C for 10 minutes. Completed PCR reactions were held at 4°C. A random subset of samples from the PCR plate were selected and checked for product and fragment length *via* agarose gel electrophoresis. 4μL of loading dye was added to 2 μL of PCR sample, which was loaded onto a 1% agarose gel, along with a marker lane ladder. Electrophoresis was performed at 100 V for 30 minutes and photographed under UV. Next, PCR product was cleaned up using the SequalPrep Normalization Plate Kit (ThermoFisher Scientific, USA), where 18 μL of PCR product was taken from the PCR plate to the normalization plate. 18 μL of SequalPrep Normalization Binding Buffer was added to the 18 μL of PCR product, and mixed, sealed, vortexed, and spun briefly. Sample was incubated at room temperature for 60 minutes. Liquid was aspirated from the wells. 50 μL of SequalPrep Wash Buffer was added, pipetted up and down twice, and aspirated immediately. Next, 20 μL of SequalPrep Elution Buffer was added, mixed by pipetting up and down five times, sealed, vortexed, and spun down. Samples were incubated at room temperature for 5 minutes. Finally, 5 μL of each sample was taken and pooled and the remaining sample product frozen for later use if necessary. Next, the pooled library was diluted to 1:1, 1:10, 1:1000, 1:2000, and 1:4000 by using molecular grade water, and two QC steps performed. Firstly, diluted libraries were prepared for use on the Agilent Bioanalyzer 2100 (Agilent, USA), by loading 9 μL of gel-dye mix followed by 5 μL of marker and 1 μL of pooled library sample into a well on a High Sensitivity DNA Analysis Kit chip (Agilent, USA). Chips were vortexed for 1 minute and run on the Agilent Bioanalyzer 2100 using the HS DNA Assay setting. The next step of quality control was to perform qPCR quantification of the pooled libraries. 10 μL reactions were prepared comprised of 6 μL of KAPA SYBR FAST Q-PCR Master Mix (Merck, Germany) and 4 μL of library pool and dilutions, or 4 μL of standards (six in total run in triplicate, representing a range of dilutions, including 20 pM, 2 pM, 0.2 pM, 0.02 pM, 0.002 pM, and 0.0002 pM). Plates were placed in a thermocycler with the following conditions: Initial activation at 95°C for 5 minutes, followed by 35 cycles of; denaturation at 95°C for 30 seconds and annealing at 60°C for 45 seconds. Library concentrations were determined by the following calculation:

$$Lib_c = Avg_p \times \frac{452}{Avg_f} \times D_f$$

Where Lib_c denotes the library concentration, Avg_p denotes the average concentration of the pool, Avg_f denotes the average fragment length determined by the BioAnalyzer, and D_f denotes the dilution factor of the library pool. Normalised pools were then created by diluting to the concentration of the least concentrated sample. A single final pool was created by adding equal volumes of each post-qPCR normalised pool, to a total volume of 40–80 μL.

For sequencing, library pools were denatured by adding 5 μL of 0.2 N NaOH to 10 μL of library in a microcentrifuge tube and incubated at room temperature for 5 minutes. 990 μL of HT1 buffer was added to the mixture to create a 4 pM denatured library. A PhiX positive control was prepared by adding 2 μL of the PhiX control, 3 μL of molecular grade water, and 5 μL of 0.2 N NaOH, which was vortexed and spun for 1 minute at 400 x *g*, and 990 μL of HT1 added to create a 4 pM PhiX control. 900 μL of 4 pM library and 100 μL of 4 pM PhiX were combined. 600 μL of the combined solution was loaded on a MiSeq V2 500 bp reagent cartridge for sequencing (Illumina, USA). Demultiplexed paired-end FASTQ files were returned from sequencing by Northumbria University for subsequent pipeline analysis.

2.5.1.3 *Shotgun Metagenome Sequencing*

Diluted template DNA extractions were submitted to Northumbria University for sequencing on the Illumina MiSeq platform. DNA samples were diluted to 0.2 ng/μL by adding molecular grade water. Ligation and tagging of reads ('tagmentation', which includes fragmentation of genome DNA) was performed on genomic DNA using the NextEra XT DNA Library Preparation Kit (Illumina, USA) following the manufacturer's protocol, which also adds Illumina-specific adaptor sequences to the ends of DNA fragments. Tagmented DNA was amplified *via* a limited-cycle PCR to add Illumina-specific index sequences – index 1 (i7) and index 2 (i5) – which enables cluster formation. Indexed and tagmented DNA was cleaned up using AMPure XP beads solution (Agencourt, UK). In lieu of DNA elution with water, resuspension buffer was added to PCR products and mixed to resuspend magnetic beads. 13 μL of supernatant containing eluted DNA was transferred. DNA amplicon fragments were then checked for product and fragment size using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA) using an Agilent High Sensitivity DNA chip. DNA was quantified using a Qubit 3.0 fluorometer. Samples were prepared for sequencing in a very similar fashion to the 16S rRNA gene Illumina sequencing,

albeit with a different final library concentration and without the preparation of a PhiX control. Indexed DNA fragments were diluted to an equimolar library mix of sample DNA with a pooled concentration of 4 nM. The sample library was denatured with 0.2 N NaOH, incubated at room temperature for 5 minutes, and HT1 buffer added. A final 20 pM denatured library was then diluted to 14 pM with additional HT1 buffer and placed on a MiSeq V3 600 bp reagent cartridge for sequencing. Demultiplexed FASTQ files were returned from Northumbria University after sequencing.

2.5.2 Sequencing of the Fungal ITS Region

2.5.2.1 *Identification and Testing of ITS2 Primers for Sequencing*

The fungal internal transcribed spacer (ITS) region is located between the small 18S subunit and the large 28S subunit of the ribosomal RNA (rRNA) gene. It is around 600 nt in length and is formed of the ITS1 subsection, the 5.8S small rRNA subunit, followed by the ITS2 subunit. The region has been used as a target marker gene for fungal taxonomy for over two decades, and though many have coalesced around this gene region for the curation of reference database, it still has several drawbacks not shared with the prokaryotic analogue. This is primarily due to the hypervariability exhibited within the ITS region between species, such as in region length itself which can vary between 500 – 700 nt (Op De Beeck *et al.*, 2014). This variation prevents reliable aligning of reads due to a lack of conserved bases throughout the entire kingdom, and, therefore, prohibits phylogenetic placement of reads into a tree and phylogenetic analysis of aligned reads. However, it was still considered valuable to perform sequencing targeting the fungal community of the samples, to identify changes in fungal community within soil and sediment cores, and to establish a protocol for fungal sequencing and analysis for future use.

Primers targeting the fungal ITS2 region of the rRNA gene were used to identify the presence of fungi within the depth samples. The ITS2 region was targeted rather than the ITS1 or whole ITS regions because it is a shorter fragment length (~ 600 nt for the entire ITS region, ~ 350 nt for the ITS2 region) (Op de Beeck *et al.*, 2014) and therefore notionally better suited for DNA sequencing platforms that preferentially target shorter reads all of a similar length. Additionally, the researchers who designed the primers used for sequencing demonstrated that targeting only the

ITS2 rather than the whole ITS region provided sufficient sequencing coverage for taxonomic classification (Toju *et al.*, 2012).

As a first test to ensure amplification of ITS fragments was possible, 'end point' PCRs were performed with the primers ITS3 (5` – GCATCGATGAAGAACGCAGC, position 2024–2043; White *et al.*, 1990) and ITS4 (5` – TCCTCCGCTTATTGATATGC, position 2390–2409; White *et al.*, 1990). PCRs were 10 μL PCR reactions, comprised of: 9 μL of MegaMix Blue PCR Master Mix, containing dNTPs, *taq* polymerase, $MgCl_2^-$, buffer, and water (MicroZone, UK); 0.25 μL of 100 pM forward and reverse primer each, and; 0.5 μL of diluted template DNA. DNA extracted from *Trichoderma reesei* was used as a procedural positive (DSMZ, Germany). All PCRs were performed on a Techne 512 thermocycler. Initial denaturation was 95°C for 10 minutes, followed by 35 cycles of 95°C for 1 minute, 55°C for 1 minute, 72°C for 1 minute, and a final elongation step of 72°C for 10 minutes. PCR amplicons were product checked by electrophoresis on a 1% agarose gel as above. For preparation for DNA sequencing, alternative primers were used that were specifically designed to be optimal for next generation sequencing platforms. The ITS3_KYO2 forward primer (Toju *et al.*, 2012) was used as the forward primer, with Golay barcodes and Torrent Adaptor A added for identification, while the original ITS4 reverse primer was used with the trP1 adaptor, with a final target fragment length of 450 nt. A series of dilutions were tested to determine the optimum DNA concentration for use with these primers and included undiluted (neat) template, 1:10, 1:20, 1:50, and 1:100 dilutions. A range of PCR cycling conditions with various annealing temperatures were also tested to determine optimum conditions, including: 58°C, 56°C, 54°C, 52°C, and 50°C, for 1 minute. It was determined that 1:20 diluted DNA and an annealing temperature of 55°C yielded the best PCR amplicon bands, as characterised by the highest band intensity and lack of multiple band fragments per sample. This check is especially pertinent when amplifying the ITS region, as it is a hypervariable region and the length can vary dramatically between species. Multiple PCR amplicon bands is likely more representative of true biological diversity within the ITS2 region, however, modern sequencing platforms are not well-equipped at coping with multiple target fragment sizes, and preferentially sequence smaller fragments over larger ones, and therefore, it was imperative to reduce the presence of multiple fragment sizes (and in turn, likely lose out on maximising species diversity) to one amplicon fragment size. Excluding

fragments that were not the specific size potentially resulted in fungal classes including the *Glomeromycota* being absent in the subsequent taxonomic assemblages.

Samples were prepared for sequencing for Ion Torrent as above. Due to the hypervariability of the region and the range of ITS2 fragment size between fungal clades, the Pippin Prep step was not performed on the amplicon pooled library prior to sequencing. Sequencing was performed internally on the Ion Torrent Personal Genome Machine using a 316 chip (ThermoFisher Scientific, USA). Sequencing produced single-end reads in a multiplexed FASTQ file.

2.6 16S rRNA Gene Sequencing Data – Pipeline Analysis

2.6.1 Development of QIIME1 Pipeline and Exploration of OTU Picking Methods

The following method development was performed regarding the use of the QIIME1 pipeline applied to Ion Torrent sequence data. At the time of sequencing, the latest stable version of QIIME available was QIIME 1.9.1 (Caporaso *et al.*, 2010a). Sequencing analysis was initially performed using a pipeline developed internally within the school of Civil Engineering and Geosciences based on QIIME 1.7 that used a series of default options for analysis that included: default quality filtering of a Phred Q score (base calling accuracy, *i.e.*, how certain is it that a base call is correct, a higher score denotes higher certainty) of 20; 'UCLUST' (Edgar, 2010) for OTU picking; only using 97% similarity for clustering of sequences; Greengenes 13.8 reference database for OTU picking, alignment, and taxonomic assignment, and; only performing open-reference OTU picking. Recently, open-reference OTU picking using UCLUST was found to be over-inflating OTU numbers, especially of *de novo* OTUs, which also has an adverse effect on alpha diversity indices (Edgar, 2017). Briefly, OTU picking is performed using a reference database (*e.g.*, Greengenes, SILVA, RDP) of a particular similarity threshold (*e.g.*, 97%, 99% – the percentage similarity sequences have to be before they are classed as identical) by matching sequenced reads against the reference database. Sequenced reads that meet the percentage similarity requirement with a sequence in the reference database are immediately classed as OTUs. Sequences that do not match are discarded in 'closed-reference' OTU picking. In open-reference, these 'failed' sequences get

clustered amongst themselves at the percentage similarity threshold, and a representative sequence comprised of the highest quality base sequence is then matched again to the reference database. Sequences that now match enter the OTU table. It is this second round of 'failures' that become known as *de novo* OTUs. Taxonomy is still assigned based on the closest highest match. Different OTU picking methods employ different ways of matching sequences both against the reference database and also when clustering 'failures' and can dramatically impact the final number of OTUs identified.

As it was not possible to change settings in this pipeline, an updated pipeline using alternative OTU picking methods and reference databases was written. As part of developing this new pipeline, alternative OTU picking methods were explored to investigate the documented 'OTU inflation' effect observed in open reference OTU picking by UCLUST.

Three OTU picking methods available in QIIME1 open reference OTU picking (*pick_open_reference_otus.py*) were compared – UCLUST (Edgar, 2010), VSEARCH (Rognes *et al.*, 2016), and SortMeRNA (Kopylova *et al.*, 2012). Additionally, OTU numbers from both 97% and 99% similarity thresholds were compared, as well as OTU numbers between open reference, and closed reference (*de novo* OTUs removed) OTU picking. Lastly, the SILVA 119 (Quast *et al.*, 2011) reference database was employed instead of Greengenes 13.8 (DeSantis *et al.*, 2006) as SILVA is still curated, maintained, and developed – as well as being the only reference database which supports Tax4Fun, as discussed later. The results of this comparison are presented below in Table 2.3.

99% similarity produced more OTUs from open reference OTU picking in all three OTU picking methods than 97% similarity did, as would be expected due to increased resolution between clusters—UCLUST and VSEARCH had an increase of 91% and 261% from 97% to 99% respectively, while SortMeRNA had an increase of 69%. However, when *de novo* OTUs were filtered to create a closed-reference OTU table, the number of 99% OTUs was lower in all three methods than 97%. UCLUST and VSEARCH showed decreases of 53% and 40% respectively, while SortMeRNA had a 21% decrease. SortMeRNA also used fewer reads during OTU picking across the three methods. This indicated that SortMeRNA is the most conservative method, as SortMeRNA had by far the fewest number of OTUs in the open reference table, and the most similar number of OTUs between both 97% and 99% and open and

closed reference OTU picking. Since the OTU-inflation issue has such adverse effects on alpha diversity indices, SortMeRNA was used for subsequent pipeline analysis. 99% was used as the similarity threshold as it more accurately distinguishes between subspecies, and while taxonomic classification is unlikely to reliably classify to this level (and in turn affect functional interpretation), it is important to resolve subtle differences within the microbial communities at the OTU level within taxonomic clades. Lastly, due to the drastically higher numbers of OTUs in the open-reference tables, which are likely exclusively the result of spurious *de novo* OTUs that have not directly matched against the reference database, it was decided that any *de novo* OTUs were likely unreliable, and so closed-reference OTUs were the main focus of analysis. Subsequent versions of the SILVA database were released after sequencing data had been obtained, and pipeline analysis was re-run with newer databases, which ranged from SILVA 123, 128, to 132 (released December 2017).

A version of the scripts written is available at https://github.com/peterleary/q2pipeline. Briefly, the pipeline executes a series of QIIME1 scripts that analyse the sequencing data, in the order of: converting the FASTQ into separate 'QUAL' and 'FASTA' files; demultiplexing, which includes filtering of low-quality sequences based on Phred Q score and removing short and long reads; identifying and removing chimeric sequences (*via* VSEARCH); open-reference OTU picking (using method chosen by user, SortMeRNA recommended); aligning of sequences using PyNAST (Caporaso *et al.*, 2009) against the SILVA core alignment file; filter alignment, *i.e.*, remove positions that are gaps in every sequence; filter reads that failed to align; create phylogenetic tree from filter aligned sequences using FastTree (Price *et al.*, 2009); filter *de novo* OTUs from the open-reference OTU table to create a second, closed-reference OTU table; perform core diversity analyses on both the open-reference and closed-reference OTU tables.

2.6.2 ITS2 Sequencing Analysis

Fungal ITS2 sequencing data was also analysed in QIIME1 in a pipeline similar to that written for 16S rRNA sequencing. However, due to the hypervariability of the ITS gene, it is not possible to reliably align the sequences and therefore not possible to derive meaningful phylogenetic placement of the reads. Due to this, no aligning, alignment filtering, or alpha diversity (as part of core diversity analyses) was

Table 2.3. Tables comparing number of unique OTUs and ASVs identified from the same 16S rRNA gene sequencing data using a variety of methods. Table 2.3A shows the numbers of OTUs that were identified, and from how many reads, by open-reference OTU picking in QIIME 1.9.1 by UCLUST, VSEARCH, and SortMeRNA OTU picking methods, using the SILVA128 reference database. OTU picking was also performed at 97% and 99% similarity thresholds for OTU clustering of reads. Table 2.3B shows the numbers of OTUs within the tables that remained after all *de novo* OTUs had been filtered out, leaving only closed-reference OTUs. Table 2.3C shows the equivalent numbers derived from ASV identification in QIIME2, which does not use a reference database and only clusters identical sequences (*i.e.*, no similarity threshold).

A. QIIME 1.9.1 – Open Reference OTU Picking						
Method	**UCLUST**		**VSEARCH**		**SortMeRNA**	
Similarity	**97%**	**99%**	**97%**	**99%**	**97%**	**99%**
OTUs	41,480	79,019	40,335	145,462	9,556	16,138
Sequences	1,808,791	1,445,111	1,835,944	1,541,849	1,633,514	961,324
B. QIIME 1.9.1 – *De novo* OTUs Filtered (Closed Reference)						
Method	**UCLUST**		**VSEARCH**		**SortMeRNA**	
Similarity	**97%**	**99%**	**97%**	**99%**	**97%**	**99%**
OTUs	8,248	3,861	7,681	4,575	8,892	7,065
Sequences	1,064,909	297,639	1,175,158	323,362	1,400,985	562,013

C. QIIME2 (Amplicon Sequence Variants)		
Method	**DADA2**	**Deblur**
ASVs	10,799	3,156
Sequences	1,597,901	87,288

performed. The UNITE (Kõljalg *et al.*, 2013) reference database at variable similarity threshold (*i.e.*, between 95-99% depending on the taxonomic clade) was used for OTU picking and taxonomic assignment. Long sequences were not filtered, but instead treated as true biological sequences, again due to the hypervariability of the ITS region, although short reads (> 200 nt) were still filtered. Lastly, only closed-reference OTUs were investigated, for reasons stated above.

2.6.3 Moving to QIIME2

In January 2018, QIIME2 was released and QIIME1 was deprecated. QIIME2 brought with it several new methods for analysing marker gene sequencing data, that were considered more conservative and reliable. It also focused on reproducibility, with the inclusion of provenance information for all data created. QIIME2 stores all data in a QIIME Zip Archive (QZA) which contains the raw data and provenance data, and these can be turned into interactive QIIME Zip visualisation (QZV) files. One fundamental difference between QIIME 1 and 2 is that there was a shift from

reference-based OTU picking by clustering sequences at a specific similarity threshold, to an entirely *de novo* clustering of sequences at 100% similarity into representative amplicon sequence variants (ASVs), with taxonomic classification being performed subsequently. These ASVs are treated as true biological sequences and are comprised of sequences clustered at 100% similarity—identical reads at the single nucleotide level—and are not to be interpreted as species or subspecies level, but merely as unique sequence variants distinguished from one another by their sequences. The two options for the identification of ASVs in QIIME2 are DADA2 (Callahan *et al.*, 2016) and Deblur (Amir *et al.*, 2017). The two methods employ different techniques, with DADA2 consistently identifying more unique ASVs than Deblur, based on similar number of reads after QC (Nearing *et al.*, 2018 [preprint]). Table 2.3 shows the number of ASVs and reads identified and used by both DADA2 and Deblur, which show DADA2 identified more ASVs and used significantly more reads—likely because Deblur uses reads only of the exact length specified by the user, whereas DADA2 allows for some flexibility in read length. This is important as the 16S rRNA gene can vary in length slightly (and the ITS region can vary greatly in length), and variances in length can represent true biological differences (Head *et al.* 1996). As DADA2 is the method favoured by the developers of QIIME2, it was used in all subsequent analysis. Secondly, ASVs now have taxonomy classified to them using a naive Bayes trained classifier, rather than an assignment method based on reference-based picking. Here, reads from the target region that was sequenced are extracted from the reference database of choice by specifying the primers used during PCR, *e.g.*, sequences from the V4-V5 region targeted during Ion Torrent sequencing are extracted by entering the primers and target fragment length. Then the extracted sequences are trained using a machine-learning algorithm (based on the RDP classifier). This method of taxonomic classification was found to have superior accuracy on mock communities than the methods used in QIIME1 (Bokulich *et al.*, 2018). Notably, the reference database still clusters reference reads at a specified similarity threshold, where 99% was chosen as the highest resolution achievable with the current methods. This means that potentially different ASVs will be assigned the same taxonomy, resulting in a (albeit minor) loss of resolution. Lastly, aligning of sequences is also performed *de novo* and not by using a reference database (as was done previously in QIIME1), using MAFFT (Katoh and Standley, 2013), again reflecting a shift away from reliance on reference-based analysis that

may adversely influence data interpretation by spurious OTU creation and unreliable taxonomic classification.

2.6.4 DADA2 and Amplicon Sequence Variants

DADA2 (Callahan *et al.*, 2016) was the sole method for ASV identification used throughout this study, though various parameters were changed between Ion Torrent and Illumina sequencing data to optimise ASV identification and subsequent taxonomic classification. Though the pipeline scripts allowed for the manual selection of sequence trimming and truncating for all reads, for the purposes of this research, the same numbers were used throughout. For Ion Torrent, all reads were trimmed at the 5` end by 20 nt, to ensure removal of any non-biological bases (*e.g.*, primer adaptors, tag sequences), and truncated at the 3` to 270 nt, leaving fragments of 250 nt that mainly covered the V4 region of the 16S rRNA. Truncation length was based on a consistent pattern observed in all Ion Torrent sequencing data that showed a step decrease in base sequence quality at around 300 nt. Illumina sequencing data was trimmed at the 5` end by 13 nt (to remove any non-biological bases) and truncated at the 3` to 240 nt for the forward reads, and 180 nt for the reverse reads – again, due to the decrease in base quality in the reads after these base numbers. Once paired reads were merged, ASVs were 227 nt in length. The other major difference between using DADA2 on Ion Torrent and Illumina sequencing data was relaxing the max expected error rate from the default of 2 to 5 for Ion Torrent. This was because Ion Torrent sequencing generally returns lower quality Phred scores and so is likely to contain more base errors. Keeping the max expected error at 2 for Ion Torrent resulted in too few sequences being used for ASV identification to be produce viable results. This parameter change is recommended in the DADA2 tutorial (https://benjjneb.github.io/dada2/tutorial.html).

2.6.5 Closed-Reference OTU Picking

Vsearch (Rognes *et al.*, 2016) was used to perform closed-reference OTU picking, by clustering the ASV representative sequences at 99% similarity, using SILVA 119 as the reference database. 99% similarity was used over 97% to keep resolution between OTUs as high as possible. The OTU table created from this step was used only in Tax4Fun.

2.6.6 Taxonomy Classification

Representative sequences of the ASVs were classified taxonomy using a Naïve Bayes trained classifier (Bokulich *et al.*, 2018). The classifier was trained on sequences from SILVA 132, with extracted reads from either the V4 region of the 16S rRNA gene and trimmed to 251 nt for Illumina sequencing, or from the V4-V5 region and trimmed to 450 nt for Ion Torrent sequencing. Taxa bar plots showing relative abundance of each taxonomic classification were produced, where ASVs were collapsed into one when classification was identical, at each level from Kingdom through to species. Relative abundances were calculated by taking the sum of the identically classified ASV abundances and dividing that by the total abundance of all ASVs within a sample, to give relative abundance as a fraction of one. Abundance values presented in this thesis represent ASV taxonomic classification abundances collapsed at genus level, *i.e.*, all ASVs that had the same classification to genus level were merged into one ASV. Not all ASVs classified down to a specific named genus however, and therefore the lowest specific named classification level is presented, but still represents ASVs merged at the genus level.

2.6.7 Sequence Alignment and Phylogenetic Trees

To perform phylogenetic analysis, sequences were first aligned *de novo* using MAFFT (Katoh and Standley, 2013), to infer homology and evolutionary distance between sequenced reads (representative sequences of the ASVs). Aligned sequences were then masked to remove highly variable positions that would adversely affect inferred homology. These masked aligned sequences were then placed into an unrooted phylogenetic tree using FastTree (Price *et al.*, 2009), and then rooted at the mid-point, determined by the longest tip-to-tip distance in the unrooted tree between two masked aligned reads. This mid-point rooted tree was then used to perform a series of alpha and beta diversity analyses.

2.6.8 Alpha and Beta Diversity Analyses

As part of the ‘core-metrics-diversity’ plugin in QIIME2, a series of alpha and beta diversity analyses were performed on data. Data were first rarefied by randomly

sub-sampling each sample to a specified sampling depth, usually determined by the sample with the lowest number of reads. Alpha diversity metrics used to compare community diversity between samples within each core was focused on: Pielou's Species Evenness; Faith's Phylogenetic Diversity (Faith PD), and; Shannon's species richness diversity indices. These three indices represented a mix of quantitative and qualitative measures of community composition. Boxplots denoting range in index values within and between arbitrary sample groups (*e.g.*, samples clustered into groups of three in order of depth) based on the three indices were produced. These were used to describe trends in alpha diversity indices as a function of increasing depth within the core analysed in this study. Group significance was performed *via* Pairwise Kruskal-Wallis ANOVA for all three alpha diversity metrics on sample groups to derive H and *p* values, detailing statistical differences in community composition within sample groups. Whisker boxplots of the three indices were produced in R using the vectors created in QIIME2.

Beta diversity within QIIME2 encompassed creating distance matrices based on the phylogenetic tree using the weighted Bray Curtis dissimilatory, and weighted and unweighted UniFrac metrics, where the matrices produced by these metrics was used to produce 3D principal coordinate analysis (PCoA) ordination plots. Group significance was performed *via* Pairwise PERMANOVA for both beta diversity metrics on sample groups to derive *p* values, detailing statistic differences in community composition between sample groups.

2.6.9 Tax4Fun

Tax4Fun (Asshauer *et al.*, 2015) is an R program that aims to construct inferred functional gene profiles by using precomputed association matrices and metabolic reference profiles to link 16S rRNA gene taxonomy to the most relevant fully sequenced genome *via* nearest-neighbour identification. Tax4Fun relies on the SILVA database and on precomputed association matrices linking SILVA taxonomy to KEGG organisms. This means that Tax4Fun limits the version of SILVA that can be used to the version used to create the precomputed matrices, which is currently SILVA 119 or 123. All default settings were used for Tax4Fun, including: fctProfiling = TRUE (uses the precomputed KEGG Ortholog reference pathways); refProfile =

Table 2.4. A list of the functional genes of interest, and their respective KEGG Ortholog numbers, that were identified using the Tax4Fun functional profiling of 16S rRNA gene sequences.

Functional Gene Target	KEGG Ortholog
Anaerobic carbon-monoxide dehydrogenase – *cooS*	K00198
Nitrite reductase – *nirK*	K00368
Nitrate reductase – *narG*	K00370
Nitrous-oxide reductase – *nosZ*	K00376
Adenylylsulfate reductase – *aprA*	K00394
Methyl coenzyme M reductase – *mcrA*	K00399
NiFe hydrogenase (cytochrome c3 hydrogenase) – *hydB*	K00437
Nitrate reductase – *napA*	K02567
Nitrogenase – *nifD*	K02586
Nitrite reductase – *nrfA*	K03385
Heterodisulfide reductase – *hdrA*	K03388
Nitric oxide reductase – *norB*	K04561
Methylene-tetrahydromethanopterin dehydrogenase – *mtdB*	K10714
Ammonia monooxygenase – *amoA*	K10944
Dissimilatory sulphite reductase – *dsrA*	K11180
Methanol dehydrogenase – *mxaF*	K14028
Acetyl CoA synthase – *acsB*	K14138
Nitrite reductase – *nirS*	K15864
Methane monooxygenase – *mmoX*	K16157

"UProC" (uses the UProC method for precomputed reference profiles); shortReadMode = TRUE (uses 100 nt reads); normCopyNo = TRUE (taxonomic profiles are normalised by the 16S rRNA gene copy number). Tax4Fun outputs an inferred functional gene profile table comprised of columns of samples and rows of inferred functional genes. Abundances of genes are presented as a fraction of 1 for each sample column and have been normalised by 16S rRNA gene copy number of

each ASV used. The output functional gene profile was saved to file, and a script was written that extracted relevant KEGG Ortholog gene abundances that comprised microbial metabolic pathways, *e.g.*, the KEGG Ortholog genes from denitrification or for acetoclastic methanogenesis, and are presented in Table 2.4. The inferred abundances of specific target genes were analysed, and where specific genes were also targeted by other methods, abundances were directly compared.

2.6.10 QIIME2 Pipeline Development

A set of scripts to automate pipeline analysis was authored to more efficiently utilise QIIME2 and the new ASV identification, taxonomic classification, and alignment methods. This set of scripts and an in-house tutorial were prepared to facilitate the use of the analysis pipeline by users that are not familiar with command-line interface. The scripts for the pipeline are a maintained and updated version available for download at https://github.com/peterleary/q2pipeline. The pipeline analyses both single-end multiplexed Ion Torrent data and paired-end demultiplexed Illumina data (with support for ITS analysis in the future). A brief overview of the pipeline is as follows: import data into compatible QZA format, where a multiplexed Ion Torrent FASTQ file is first split into separate reads and barcode FASTQ files, as per the Earth Microbiome Project (EMP) protocol, and demultiplexed Illumina FASTQ files are simply converted into a single QZA file; demultiplex the sequences based on barcode [Ion Torrent only]; use DADA2 to filter, trim and truncate reads (based on user-input parameters), learn the error rate for each nucleotide base using a parametric error model, dereplicate sequences, infer ASVs from reads, merge paired reads [Illumina only], and construct the sequence table (analogous to an OTU table); create a closed-reference OTU table using the SILVA119 database and VSEARCH for use with Tax4Fun; perform Tax4Fun on the closed-reference OTU table; classify taxonomy and create taxa bar plots; align sequences; mask aligned sequences; create a phylogenetic tree using FastTree (Price *et al.*, 2010); perform Gneiss differential abundance if applicable (Morton et al., 2017), and; perform core metric diversity to retrieve alpha and beta diversity analyses. Lastly, the pipeline organises the main outputs into one folder, exporting key QZA and QZV files, such as the taxa bar plots, alpha and beta diversity visualisations, and ASV tables and representative sequences file.

2.7 Shotgun Metagenome Sequencing Analysis

2.7.1 GraftM

GraftM (Boyd *et al.*, 2018) is a bioinformatics tool written in Python that uses precompiled functional gene packages containing Hidden Markov models (HMMs) to identify functional gene families within metagenome sequencing data. It then classifies these identified functional gene reads into a preconstructed gene tree (from the gene package), and uses KronaTools (Ondov *et al.*, 2011) to demonstrate phylogenetic placement of functional genes in a metagenome sequencing dataset. GraftM utilises a range of other programs, including PPLACER (Matsen *et al.*, 2010) for phylogenetic placement of reads into trees, FastTree (Price *et al.*, 2009) also for phylogenetic tree placement, MAFFT (Katoh and Standley, 2013) for alignment of reads, FXTRACT (https://github.com/ctSkennerton/fxtract) for extraction of reads from FASTQ and FASTA files, HMMER (v3.1b2, http://hmmer.org) for creation of HMM profiles, OrfM (Woodcroft *et al.*, 2016) for open reading frame calling, and KronaTools (Ondov *et al.*, 2011) for visualisation of phylogenetic placement of reads. Additionally, DIAMOND (Buchfink *et al.*, 2015) is also used by GraftM to align protein sequences, but this tool was not used in this project. GraftM is capable of parallelised placement of functional gene reads into phylogenetic trees from marker gene and shotgun metagenome sequencing data and was demonstrated by the authors to be fast and accurate (Boyd *et al.*, 2018). Results discussed from using GraftM and a functional gene package on shotgun metagenome sequence dataset include relevant phylogenetic changes of functional genes between samples (especially in the context of depth relevant to the cores analysed in this study), as well as the number of reads identified by GraftM for each functional gene within each sample.

GraftM precompiled functional gene packages are available at https://github.com/geronimp/graftM_gpkgs and include many relevant functional gene targets in soils and sediments, as well as a package for identification, alignment, and phylogenetic placement of 16S rRNA gene fragments based on the Greengenes 13.8 reference database. A list of GraftM packages downloaded and used in this project is included in Table 2.5. In addition, three GraftM packages were made using the ‘Create’ function in GraftM. The first was a package based on the SILVA 132 reference database for the 16S rRNA gene. A copy of the SILVA 132 files for QIIME were downloaded from the SILVA website (https://www.arb-silva.de/fileadmin/silva_databases/qiime/Silva_132_release.zip). The representative

sequences file and aligned representative sequences file, the raw taxonomy file, and the tree file for the 97% clustered OTUs were used in creating the GraftM package. The second package created was for the *dsrA* gene. To create this, the functional gene repository FunGene (Fish *et al.*, 2013) was used to download all the sequences in the *dsrA* gene collection, a total of 8,761 sequences (as of 16th February 2018) with a modal length of 369 nt. A FASTA file of unaligned sequences was downloaded. Taxonomic information was separated to create a taxonomy file. Sequences were clustered into OTUs at 90% similarity to dereplicate the sequence file, giving a total of 2,565 OTUs. This was performed using the *pick_otus.py* function in QIIME 1.9.1 using the VSEARCH method and 90% similarity threshold, and all other options as default. A representative sequence for each OTU was created using the *pick_rep_set.py* function, using the OTU map created from *pick_otus.py* and the original unaligned fasta file, with all other options as default. This representative sequence FASTA file was loaded into QIIME2, where it was aligned, the alignment masked, a phylogenetic tree created, and then rooted, as described by the methods in section *2.6.7.* The masked alignment FASTA and mid-point rooted Newick tree files, along with the raw unaligned FASTA and parsed taxonomy files, were used in GraftM to create a package for the *dsrA* gene. The third package was for the *amoA* gene and was comprised of genes downloaded from FunGene functional gene repository for both ammonia oxidising bacteria and archaea (AOB and AOA respectively). Since both these gene repositories contained more than 40,000 sequences, and FunGene limits the number of sequences able to be downloaded at once to 10,000, the AOB and AOA repositories were filtered based on their score until a selection of > 10,000 sequence. For AOB, this was to a score of 327 and for AOA, the score used was 458, to give 9,840 and 9,783 sequences for AOB and AOA respectively. Sequences were clustered into OTUs at 97%. Briefly, to create a GraftM package, an alignment with an HMM is produced using HMMER, this alignment is built into a tree using FastTree, and the package is compiled using Taxtastic (https://github.com/fhcrc/taxtastic). Filtering of short or unaligned reads is included between these steps.

2.8 Statistical Analysis and Graphics in R

A range of statistical analytical methods were used to describe trends regarding microbial taxonomic and functional gene abundances. Statistical analyses

Table 2.5. List of genes targeted by the GraftM program. All packages were downloaded from https://data.ace.uq.edu.au/public/graftm/7/ except genes with * which were created using the Create tool within GraftM, as described in the main text.

Gene	Description	KEGG Ortholog
accC	Acetyl-CoA carboxylase	K01961
aclA	ATP-citrate lyase alpha subunit	K15230
acsB	Acetyl-CoA synthase	K14138
****amoA***	Ammonia monooxygenase subunit A	K10944
aprA	Adenylylsulfate reductase subunit A	K00394
****dsrA***	Dissimilatory sulphite reductase subunit A	K11180
16S rRNA	Greengenes 13.8 16S rRNA gene	K01977
mcrA	Methyl Coenzyme M reductase subunit A	K00399
mxaF/xoxF	Methanol dehydrogenase subunit 1	K14028
narG	Nitrate reductase alpha subunit	K00370
hydB	NiFe hydrogenase large subunit	K00437
nifH	Nitrogenase iron protein	K02588
nirK	Nitrite reductase (NO-forming)	K00368
nosZ	Nitrous-oxide reductase	K00376
sat	Sulfate adenylyltransferase	K00958
****soxB***	Sulfur oxidising protein	K17224

were performed using R 3.5.1 (R Core Team, 2018), QIIME2, and PRIMER V6. Statistical calculations were presented was *p* values, where *p* values close to zero represent statistically significant results. A threshold *p* value of < 0.05 was indicative of a significant result, where a *p* value < 0.05 indicates that there was < 5% risk of wrongly determining a significant result when it is in fact not (a false positive, or type II error).

Normality testing was performed on all datasets in R 3.5.1 visually by the generation of density plots and Q-Q plots, and statistically via Shapiro-Wilk normality test (where p value > 0.05 implies the data is not statistically different from normal distribution), allowing for univariate statistical methods such as one-way analysis of variance to be used on the sample groups.

Samples were put into arbitrary groups to enable intra-core comparison of measured values. The Needle's Eye depth cores were put into four groups made up of: 1 – 10 cm (Group 1); 11 – 20 cm (Group 2); 21 – 30 cm (Group 3), and; 31 – 41 cm (Group 4). Ravenglass was divided into 1 – 5 cm, 6 – 10 cm, 12 – 20 cm, and 22 – 28 cm. Mud Patch was divided into 1 – 5 cm, 6 – 10 cm, 12 – 20 cm, 22 – 30 cm, and 32 – 36 cm. Loch Etive cores were grouped by the resolution at which they were

sampled, *i.e.*, all the samples from one core that were sampled at 0.5 cm resolution formed one group, the samples at 1 cm resolution the next group, and so forth.

One-way analysis of variance (AOV) was performed across all sample groups from within each individual core, for DNA, total organic carbon (TOC), alpha diversity index values, ASV, Tax4Fun inferred functional gene abundances, and geochemical data where possible, to derive *p* values to discern whether there were statistical differences between abundances across all samples. *Post hoc* Tukey's honest significant difference (HSD) test was also performed pairwise between all pairs of sample groups within each individual core, to determine which sample group pairs had statistical differences in abundances.

The four genes targeted by 'end point' PCR and the corresponding inferred abundances of those genes from Tax4Fun, and extracted DNA and TOC were compared by way of XY scatterplot. Spearman correlation was performed to determine the strength of the association between abundances of the same gene between the two methods by ρ and *p* values generated.

2.8.1 Beta Diversity Ordination Plots in R

In addition to the PCoA plots created by QIIME2, additional ordination methods were used to compare ASV assemblages between samples both between two cores from a site and within individual cores. This was in an attempt to see if various ordination methods reconciled in the interpretation of community assemblage similarity derived from the methods. Ordination of ASV taxonomic classification collapsed at different levels (class, order, family) was performed where ASV assemblages were completely dissimilar, to allow comparison of taxonomic classification similarity between samples from two cores (Ravenglass and Mud Patch, and Loch Etive). The methods were non-metric multidimensional scaling (nMDS) plots based on Bray Curtis and Jaccard dissimilarity metrics, and detrended correspondence analysis (DCA). Ordination was performed using the 'vegan' package in R 3.5.1, specifically the 'metaMDS' function for nMDS and 'decorana' for DCA. Hierarchical clustering of samples based on centroid similarity was performed and overlaid on the nMDS plots to show samples that clustered based on community similarity. Bray Curtis, weighted, and unweighted UniFrac dissimilarity vectors from QIIME2 were imported into R using 'qiime2R' (v0.12,

https://github.com/jbisanz/qiime2R) and 2D PCoA ordination plots created using the 'MicrobeR' (v0.3, https://jbisanz.github.io/MicrobeR/) package.

Taxonomic classifications were collapsed at a range of classification levels as described by merging ASV tables that contained taxonomic classification data using the *merge_otu_tables.py* command in QIIME1 (version 1.9.1) followed by the *summarize_taxa.py* command (with the -L flag set to the desired level to collapse taxonomic classifications). The resulting table was then used to generate nMDS and DCA plots.

2.8.2 Environmental Fitting

Non-metric multidimensional scaling (nMDS) plots based on Bray Curtis dissimilarity and detrended correspondence analysis (DCA) were produced as above, using the 'vegan' package in R 3.5.1. Additionally, geochemical data was superimposed onto the nMDS and DCA using the 'envfit' fitting of environmental variable function in 'vegan', where the projections of points onto vectors have maximum correlation with corresponding environmental (geochemical) variables.

'Envfit' fits vectors of continuous variables and centroids of levels of class variables. The arrow shows the direction of the (increasing) gradient, and the length of the arrow is proportional to the correlation between the variable and the ordination (Oksanen, https://cran.r-project.org/web/packages/vegan/vignettes/intro-vegan.pdf).

Essentially, geochemical variables were correlated with nMDS and DCA vectors, and geochemical variables that have a statistically significant correlation (*i.e.*, p value < 0.05) are fitted onto the nMDS and DCA as lines. The direction and relative length of the line is determined by the correlation of to each nMDS and DCA vector. This is a qualitative way of correlating abundance of geochemical variables with samples.

2.9 Geochemical Analysis

All geochemical analysis was conducted by the University of Manchester.

The contaminated core was prepared for solid phase elemental analysis by X-ray fluorescence (XRF). Prior to analysis the sections were freeze-dried, homogenised, and ground to a talc-like consistency in an agate ball mill. The

concentration of Al, Fe, K, Mg, Mn, and P in the bulk sediment in samples was determined using a Thermo ARL 9400 XRF Sequential X-ray Fluorescence Spectrometer, where samples were pressed into a pellet with wax binder and analysed, with data oxide normalised for C content.

Water content and both organic matter and total carbon content were determined by sequential loss on ignition. Here, 2 g of field-moist soil or sediment from each depth was heated to 105°C to determine water content, 375°C for organic matter content, and 950°C for total carbon. Samples were cooled in a desiccator and reweighed after each temperature. For ferrozine analysis, sub-samples of cores sectioned under N_2 were used for measurement of the solid phase 0.5 N HCl extractable $Fe^{2+/3+}$ ratio (Lovley and Phillips, 1987). Additional sub-samples of this core were used for X-ray absorption spectroscopy (XAS).

Porewater samples (10 mL) were then extracted with the initial 1 mL discarded. Samples were passed through a poisoned 0.22 µm cellulose filter (to avoid microbial contamination) into two sample bottles, one with 10 µL of concentrated Aristar grade HNO_3 for metal analysis and a second with Ar purged headspace for anions. Dried and ground plant tissue (0.1 g dry weight) was incubated in 5 ml of 70% ultrapure grade HNO3 at 140 ºC for 3 h. Cooled samples were diluted by addition of MilliQ H_2O to a final volume of 10 ml then filtered through a 0.45 µm Millipore MCE membrane before a final dilution with MilliQ H_2O to 10 ml. Aqueous U and Fe concentrations were subsequently measured on an Agilent 7500cx ICP-MS with inline calibration standards and major anions were measured by ion chromatography on a Dionex ICS5000.

2.9.1 Needle's Eye Contaminated Core X-Ray Absorption Spectroscopy

Select samples (1, 5, 10, 15, 20, 23, 25, 27, 30 and 35 cm) were analysed by U L_3-edge (17166 eV) XAS to determine variation in U oxidation state and coordination environment with depth. Spectra were collected on beamline B18 at the Diamond Light Source, UK and the MARS beamline at Synchrotron Soleil, France. Both B18 and MARS use bending magnet sources to generate X-rays. On both beamlines the edge energy was selected by a double crystal Si 111 monochromator. In both cases samples were analysed in a liquid N_2 cryostat to improve data quality and reduce beam damage. Spectra were collected in fluorescence mode using a 36-

element solid state Ge detector at B18 and a 12-element solid state Ge detector at MARS. Extended X-ray absorption fine structure (EXAFS) spectra were also collected up to 14 Å-1 in k space from select samples (5, 15 and 30 cm). A sub-set of samples (10, 20, 25, 30 and 35 cm) were also analysed by high energy resolution fluorescence detector X-ray absorption near edge structure (HERFD-XANES) spectroscopy on the U M_4-edge (3728 eV) at beamline ID26 at the European Synchrotron Radiation Facility, France. The X-rays at ID26 are generated by an undulation type insertion device. The beam energy is selected using a double crystal Si 111 monochromator and focused onto the sample. Spectra were collected using a Johann-style X-ray emission spectrometer (Roberts *et al.*, 2017).

Linear combination fitting of the U L_3-edge XANES was performed using Athena (Ravel and Newville, 2005) with end-member standards of UO_2 and UO_3. The E0 and white line position of each sample was determined, and goodness of fit was statistically evaluated. The HERFD-XANES spectra were analysed using the Iterative Transformation Factor Analysis (ITFA) software (Roßberg *et al.*, 2003). Here the relative concentration of each oxidation state was determined by fitting with standards of UO_2, UO_3, and U_3O_8. The collected EXAFS spectra were analysed using Artemis running FEFF v6.0 (Ravel and Newville, 2005). Shell by shell fitting was performed in R-space, refining the coordination number (as a function of the amplitude factor (S02), interatomic distance and Debye-Waller factor for each shell. Best fit was determined by interrogation of the fitting statistics, specifically reduction of the reduced chi-squared and the R factor. The significance of each shell was evaluated using the f-test (Downward *et al.*, 2007).

2.9.2 Ravenglass and Mud Patch Organic Matter, Porewater and Radionuclide Geochemistry

Total organic matter in the Ravenglass and Mud Patch sediments was determined *via* Sequential Loss on Ignition. Homogenised sediment (~ 5.5 g) from each sample was oven dried at 105°C for 24 hours, then heated at 375°C for 19 hours using a ramp rate of 1.0°C min^{-1}, and finally ashed at 950°C for 10 hours using a ramp rate of 10°C min^{-1}.

The concentration of Fe and Mn in the bulk sediment was determined using an Axios Thermo ARL 9400 XRF Sequential X-ray Fluorescence Spectrometer. In brief,

ground sediment (12 g) was added to powdered was binder (3 g) and homogenised (7 minutes at 350 rpm) for 2 minutes. The resulting pressed discs were analysed, and the data was oxide normalised and corrected for LOI.

The acidified porewater (200 µL) from each core was diluted with 2 % HNO3 (4.8 mL) and analysed for Fe and Mn concentrations on a Perkin Elmer Optima 5300 Dual View ICP-AES. A range of certified standards (0.01 - 5 ppm) were used for both analytes and produced a calibration line of $R^2 = 0.999$ (Fe) and $R^2 = 0.999$ (Mn).

Porewater sulfate concentrations were determined by ion chromatography using a Dionex ICS5000 chromatograph (Thermo Scientific, Sunnyvale, USA).

The activity of ^{241}Am and ^{137}Cs was determined by counting the pressed disc sample (12 g) from XRF analysis on the Ge(Li) Canberra γ-detector for 24 hours. The detector efficiency was calibrated using matrix-matched standards (of the same geometry) spiked with certified standard solutions of ^{241}Am and ^{137}Cs of known activity (Amersham International). Gamma spectra were also corrected for background. Plutonium (^{238}Pu, $^{239,240}Pu$) activity was determined using a Canberra 7401 α-spectrometer and ^{238}Pu activity using a 1220 Quantulus Ultra low-level liquid scintillation spectrometer.

2.10 References

Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Xu ZZ, Knightley EP, Thompson LR, Hyde ER, Gonzalez A, Knight R (2017) Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems, 2(2): e00191-16; https://doi.org/10.1128/mSystems.00191-16

Asshauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomics 16S rRNA data. Bioinformatics, 31(7): 2882–2884

Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso JG (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome, 6(90): https://doi.org/10.1186/s40168-018-0470-z

Boyd JA, Woodcroft BJ, Tyson GE (2018) GraftM: a tool for scalable, phylogenetically informed classification of genes within metagenomes. Nucleic Acids Research, gky174, https://doi.org/10.1093/nar/gky174

Brosius J, Palmer ML, Kennedy PJ, Noller HF (1978) Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. Proceedings of the

National Academy of Sciences of the United States of America, 75 (10): 4801-4805

Buchfink B, Zie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nature Methods, 12: 59–60

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 13: 581-583

Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R (2009) PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics, 26(2): 266–267

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Pena A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7(5): 335-336

Clarke KR, Gorley R (2009) PRIMER: Getting started with v6. Plymouth.

Cummings DS, Snoeyenbos-West OL, Newby DT, Niggemyer AM, Lovley DR, Achenbach LA, Rosenzweig RF (2003) Diversity of Geobacteraceae species inhabiting metal-polluted freshwater lake sediments ascertained by 16S rDNA analyses. Microbial Ecology, 46(2): 257–269

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Applied and Environmental Microbiology, 72(7): 5069–5072

Downward L, Booth CH, Lukens WW, Bridges F. (2007) A Variation of the F-Test for Determining Statistical Relevance of Particular Parameters in EXAFS Fits. AIP Conference Proceedings 882(1): 129-131

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26(19): 2460-1. https://doi.org/10.1093/bioinformatics/btq461

Edgar RC (2017) Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. PeerJ, 5: e3889. https://doi.org10.7717/peerj.3889

Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, Cole JR (2013) FunGene: the functional gene pipeline and repository. Frontiers in Microbiology, 4: https://doi.org/10.3389/fmicb.2013.00291

Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. Proceedings of the National Academy of Sciences of the United States of America, 102(41): 14683-14688

Gardes M, Bruns TD (1993) ITS primers with enhances specificity for basidiomycetes – application to the identification of mycorrhizae and rusts. Molecular Ecology, 2(2): 113–118

Guan J, Xia LP, Wang LY, Liu JF, Gu JD, Mu BZ (2013) Diversity and distribution of sulfate-reducing bacteria in four petroleum reservoirs detected by using 16S rRNA and *dsrAB* genes. International Biodeterioration and Biodegradation, 76: 58–66

Hamady M, Walker JJ, Kirk Harris J, Gold NJ, Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods, 5: 235–237

Head I, Gray ND, Clarke KJ, Pickup RW, Gwynfryn Jones J (1996) The phylogenetic position and ultrastructure of the uncultured bacterium *Achromatium oxaliferum*. Microbiology, 142: 2341–2345

Jarvis AP, Davis JE, Gray ND, Orme PHA, Gandy CJ (2014) Mitigation of pollution from abandoned metal mines. Environment Agency Science Report – SC090024/R3

Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution, 30(4): 772–780

Klein M, Freidrich M, Roger AJ, Hugenholtz P, Fishbain S, Abicht H, Blackall LL, Stahl DA, Wagner M (2001) Multiple Lateral Transfers of Dissimilatory Sulfite Reductase Genes between Major Lineages of Sulfate-Reducing Prokaryotes. Journal of Bacteriology, 183(20): 6028–6035

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Põldmaa K, Saag L, Saar I, Schüßler A, Scott JA, Senés C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiß M, Larsson K-H (2013) Towards a unified paradigm for sequence-based identification of Fungi. Molecular Ecology, DOI: 10.1111/mec.12481

Kopylova E, Noé L, Touzet H (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics, 28(24): 3211–3217

Law GT, Shimmield TM, Shimmield GB, Cowie GL, Breuer ER, Harvey MS (2009) Manganese, iron, and sulphur cycling on the Pakistan margin. Deep Sea Research Part II, 56: 305–323

Lovley DR, Phillips EJP (1987) Rapid assay for microbially reducible ferric iron in aquatic sediments. Applied Environmental Microbiology, 53(7): 1536-1540

Lovely DR, Giovannoni SJ, White DC, Champine JE, Phillips EJP, Gorby YA, Goodwin S (1993) *Geobacter metallireducens* gen. nov. sp. nov., a microorganism capable of coupling the complete oxidation of organic

compounds to the reduction of iron and other metals. Archive of Microbiology, 159(4): 336–344

Luton PE, Wayne JM, Sharp RJ, Riley PW (2002) The *mcrA* gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. Microbiology, 148: 3521–3530

Matsen FA, Kodner RB, Armbrust EV (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformatics 11(538): https://doi.org/10.1186/1471-2105-11-538

McCann CM (2012) The use of natural Mn oxide-containing wastes as a contaminated land remediation strategy and their effects on soil microbial functioning. PhD thesis, Newcastle University, viewed 1 June 2018, https://theses.ncl.ac.uk/dspace/bitstream/10443/1896/1/McCann%2012%20%2812mnth%29.pdf

McTavish H, Fuchs JA, Hooper AB (1993) Sequence of the gene coding for ammonia monooxygenase in Nitrosomonas europaea. Journal of Bacteriology, 175(8): 2436-2444.

Morris R, Schauer-Gimenez A, Bhattad U, Kearney C, Struble CA, Zitomer D, Maki JS (2014) Methy coenzyme M reductase (*mcrA*) gene abundance correlates with activity measurements of methanogenic H_2/CO_2-enriched anaerobic biomass. Microbial Biotechnology, 7(1): https://doi.org/10.1111/1751-7915.12094

Morton JT, Sanders J, Quinn RA, McDonald D, Gonzalez A, Vázquez-Baeza Y, Navas-Molina JA, Son SJ, Metcalf JL, Hyde ER, Lladser M, Dorrestein PC, Knight R (2017) Balance Trees Reveal Microbial Niche Differentiation. mSystems, 2:e00162-16. https://doi.org/10.1128/mSystems.00162-16.

Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology, 59(3): 695–700

Nearing JT, Douglas GM, Comeau AM, Langille MGI (2018) Denoising the Denoisers: An independent evaluation of microbiome sequence error-correction methods. PeerJ. Preprint: Date accessed May 28th 2018.

Ondov BD, Bergman NH, Phillippy AM (2011) Interactive metagenomic visualization in a Web browser. BMC Bioinformatics, 12(1): 385–394

Op de Beeck M, Lievens B, Busschaert P, Declerck S, Vangronsveld J, Colpaert JV (2014) Comparison and Validation of Some ITS Primer Paires Useful for Fungal Metabarcoding Studies. PLoS ONE, 9(6): e97629. https://doi.org/10.1371/journal.pone.0097629

Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE 5(3): e9490. https://doi.org/10.1371/journal.pone.0009490

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41(D1): 590–596

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ravel B, Newville M (2005) ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. Journal of Synchrotron Radiation, 12(4): 537-541

Roberts HE, Morris K, Law GTW, Mosselmans JFW, Bots P, Kvashnina K, Shaw S (2017). Uranium(V) Incorporation Mechanisms and Stability in Fe(II)/Fe(III) (oxyhydr)Oxides. Environmental Science & Technology Letters

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584 https://doi.org/10.7717/peerj.2584

Roßberg A, Reich T, Bernhard G (2003) Complexation of uranium(VI) with protocatechuic acid—application of iterative transformation factor analysis to EXAFS spectroscopy. Analytical and Bioanalytical Chemistry 376(5): 631-638

Stephen JR, Chang YJ, Macnaughton SJ, Kowalchuk GA, Leung KT, Flemming CA, White DC (1999) Effect of toxic metals on indigenous soil β-subgroup proteobacterium ammonia oxidizer community structure and protection against toxicity by inoculated metal-resistant bacteria. Applied and Environmental Microbiology, 65(1): 95-101

Throbäck IN, Enwall K, Jarvis A, Hallin S (2004) Reassessing PCR primers targeting nirS, nirK and nosZ genes for community surveys of denitrifying bacteria with DGGE. FEMS Microbiology Ecology, 49(3): 401-417

Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-Coverage ITS Primers for the DNA-Based Identification of Ascomycetes and Basidiomycetes in Environmental Samples. PLoS ONE, 7(7): e40863. https://doi.org/10.1371/journal.pone.0040863

Woodcroft BJ, Boyd JA, Tyson GW (2016) OrfM: a fast open reading frame predictor for metagenomic data. Bioinformatics, 321(7): 2702–2703, https://doi.org/10.1093/bioinformatics/btw241

Chapter 3. Microbial Community Analysis by Next Generation Sequencing for Biogeochemical Functional Profile Inferences in a Uranium Contaminated Soil

3.1 Introduction

Needle's Eye is a site located ~ 17 miles south of the town of Dumfries in south west Scotland, UK (54'889" N, 6'691" W). The site has been well characterised from a hydrological and geochemical perspective due to the presence of an exposed vein of pitchblende (UO_2) embedded within a fracture within the granite cliff face. The flow of groundwater through the fractured granite interacts with the pitchblende and transports uranium (U) into the nearby Southwick Merse and Solway Firth estuary (Hooker, 1991; MacKenzie *et al.*, 1991; Jamet *et al.*, 1993). The flow of groundwater over the pitchblende vein enriches the water with dissolved U^{6+} but before reaching the estuary, becomes retained in a highly organic-rich bog, where as much as 90% of the U^{6+} is retained in addition to other heavy metals including zinc (Zn) and manganese (Mn) (MacKenzie *et al.*, 1991). The accumulation of U^{6+} in this bog is anticipated to be highly localised, and U levels were predicted to be significantly lower elsewhere in the Needle's Eye area not impacted by U enriched groundwaters following a site reconnaissance prior to this study. For example, the other regions of the cliff face are not thought to contain similar veins of exposed pitchblende, and therefore other organic-rich bog areas intercepting ground water flow from then would not contain dissolved U. This was demonstrated to be the case as part of this project (see below).

U is a redox-active actinide that is found in natural environments from both geogenic and anthropogenic sources. Though naturally occurring, U poses a risk to health due to heavy-metal toxicity and radioactivity. Under natural conditions, U is generally stable in two chemical states, primarily determined by the surrounding environmental redox (E_h) and pH conditions: Oxidised U^{6+} is generally thought soluble and mobile, and; reduced U^{4+} is thought insoluble and precipitates as immobile compounds with other elements such as oxygen (O) to form uraninite (UO_2) (Choppin and Wong, 1998; Markich, 2002). The presence of other compounds such as organic matter, calcium, carbonate, and dissolved organic ligands such as humic substances can complicate the behaviour of U in natural environments (Kopylova *et al.*, 2015). Generally, in oxidised and aerobic environments of circumneutral pH, U will persist in the oxidised, soluble, and mobile U^{6+} form and in more reducing and anaerobic conditions, U^{6+} may be reduced to the immobile U^{4+} form (Runde, 2000).

This was shown to be a microbially-mediated process, whereby a range of microbial taxonomic groups utilise U as a terminal electron acceptor (TEA) during anaerobic respiration (Lovley *et al.*, 1991; Lovley *et al.*, 1993; Gorby and Lovley, 1992).

In natural environments, bacteria and archaea couple the reduction of a wide range of terminal electron acceptors (TEAs) to the oxidation of both organic matter, or reduced inorganic compounds as part of respiration coupled to oxidative phosphorylation. In sedimentary systems controlled by vertical diffusion of oxygen from the surface, stratification of redox processes is normally observed in order of thermodynamic favourability of TEAs, *i.e.*, from the TEA that produces the most energy to the least (Thauer *et al.*, 1977). Table 1.1 shows the order and redox potential of common respiratory processes in natural environments.

The process of U bioreduction in natural environments is usually dependent on the prior exhaustion of other more thermodynamically-favourable TEAs such as Oxygen, manganese and nitrate and other denitrification intermediates present in the system (Konhauser *et al.*, 2002). Thermodynamically, U reduction is similar thermodynamically to Fe reduction, and the two processes typically co-occur (Finneran *et al.*, 2002; Elias *et al.*, 2003).

Constructing profiles of microbial taxonomic and functional assemblages as a function of depth from a natural environment such as Needle's Eye presents an alternative approach to measuring the geochemical properties of such a site. Geochemical parameters such as E_h and pH can change on small spatial and temporal scales, whereas DNA can persist in soils and sediments on a time scale of weeks to months (Nielsen *et al.*, 2007).

The patterns of key functional target microbial taxa and gene distributions as a function of depth could be used to make direct inferences of the geochemical conditions and biogeochemical cycling of a site in the medium term, as consistent, progressive abundance patterns which vary incrementally and progressively with depth are likely to be a direct reflection of prevailing geochemical conditions either by exerting a selection pressure on the microbial community assemblage or as a consequence of the activities of these organisms. This approach could be used to predict the redox species and long-term fate of radionuclide contaminants based on microbial community assemblages and inferred function.

Here, a range of DNA-molecular microbial community analysis tools were used to construct profiles of microbial taxonomic and functional assemblages from two depth cores collected from the Needle's Eye natural analogue site – one from the contaminated bog, and one from a visually similar bog that had lower levels of radionuclide contamination. From these microbial profiles, inferences and descriptions of the prevailing geochemical conditions and biogeochemical cycling of a range of compounds, including uranium, were made. These inferences were compared with directly measured geochemical analysis performed by XRF, ICP-MS, and HERFD-XANES X-Ray absorption spectrometry and the results interpreted in the context to the Needle's eye physical setting.

3.2 Methods

Full methods are described in the methods and materials chapter (Chapter 2) of this thesis. A brief description of the methods used in this particular chapter are provided here.

3.2.1 Sample Collection

Two soil depth cores were collected in December 2014 and in December 2015 from two separate bogs located in the Needle's Eye site (Dumfries, UK) (54°53'15" N, 3°41'35" W). One core, as a representative, was taken from the centre of the contaminated bog, while the other control bog is located approximately 200 metres west of that location. See Chapter 2.1.1.

3.2.2 DNA Extraction and DNA Quantification

DNA was extracted from each core separately, but the method and kits used were identical. DNA was extracted from ~ 500 mg of soil for each sample, and the FastDNA Spin Kit for Soil (MP Biomedicals, USA) was used as per manufacturer's instructions. For the purposes of performing statistical analyses on the data generated on samples within each core, samples from the two cores were divided into four arbitrary groups. Sample depths 1 – 10 cm formed sample depth group 1; 11 – 20 cm formed group 2; 21 – 30 cm formed group 3, and; 31 – 41 cm formed group 4. One-way analysis of variance was performed to determine whether there

was a statistical difference in extracted DNA volumes across the sample depth groups and *post hoc* Tukey HSD was performed between individual sample groups, in R 3.5.4. See Chapter 2.2.

3.2.3 'End Point' PCR of Functional Genes

A series of primer pairs were used to target a range of functional, redox-related genes (Table 3.2). 'End point' PCR amplifications were used to provide a qualitative indication of target gene abundances within the two cores as a function of depth. Target genes were chosen to encompass the range of catabolic functional processes indicative of a variety of redox conditions. Full details of target genes and procedural positives are in the methods chapter (Tables 2.1 and 2.2). See Chapter 2.3.

3.2.4 DNA Sequencing

3.2.4.1 *16S rRNA Marker Gene Sequencing*

Initially, the contaminated core was sequencing using the Ion Torrent platform using the same protocol that was performed on the control core. However, this sequencing run produced low-quality sequencing with significantly fewer reads than would normally be anticipated. The Illumina platform became available at a much later date, and it was decided that it would be beneficial to re-sequence samples from the contaminated core. Data from the original Ion Torrent sequencing run of the contaminated core is not presented.

3.2.4.2 *Illumina Sequencing*

A total of 38 out of the 42 (including the procedural blank) diluted template DNA extractions from the contaminated core were submitted for sequencing on the Illumina MiSeq platform (Illumina, USA) at the NUOMICs centre at Northumbria University. Samples were prepped for sequencing by Northumbria University. A demultiplexed library of paired-end FASTQ files was returned from Northumbria University. See Chapter 2.5.1.2.

Table 3.1. List of functional and marker genes targeted by 'end point' PCR, description of the gene function, name of the primers, sequences of the primers from 5` to 3`, the annealing site where available, the target fragment length, the PCR cycling conditions, and the source of the primers and also PCR cycle conditions where appropriate.

Gene	Primer Name	Sequence (5' -> 3')	Annealing Site	Fragment Length	Cycle Details	Reference
AOB *amoA*	amoA-1f	GGGGHTTYTACTGGTGGT	332-349[1]	530	94C for 10 min 94C for 1 m x 42 57C for 1 m x 42 72C for 1 m x 42 72C for 15 m	Stephen *et al.* (1999) McCann (2012)
	amoA-2r	CCCCTCKGSAAAGCCTTCTTC	802-822[1]			
AOA *amoA*	amoA-1f	STAATGGTCTGGCTTAGACG	N/A[4]	635	95C for 10 m 94C for 45 s x 30 55C for 1 m x 30 72C for 1 m x 30 72C for 15 m	Francis *et al.* (2005) McCann (2012)
	amoA-2f	GCGGCCATCCATCTGTATGT	N/A[4]			
nirK	FlaCu	ATCATGGTSCTGCCGCG	568-584[2]	512	94C for 2 m 94C for 30 s x 30 57C for 1 m x 30 72C for 1 m x 30 72C for 10 m	Throbäck *et al.* (2004) McCann (2012)
	R3cu	GCCTCGATCAGRTTGTGGTT	1021-1040[2]			
nirS	Cd3af	GTSAACGTSAAGGARACSGG	916-935[3]	465	94C for 2 m 94C for 30 s x 35 57C for 1 m x 35 72C for 1 m x 35 72C for 10 m	Throbäck *et al.* (2004) McCann (2012)
	R3cd	GASTTCGGRTGSGTCTTGA	1322-1341[3]			
dsrA	DSR1F	ACSCAYTGGAARCACG	N/A[4]	c. 1900	94C for 1 m 94C for 40 s x 30 54C for 45 s x 30 72C for 2 m x 30 72C for 10 m	Klein *et al.* (2001) Guan *et al.* (2013)
	DSR4R	GTGTARCAGTTDCCRCA	N/A[4]			
mcrA	ML-f/mcrF	GGTGGTGTMGGATTCACACA--RTAYGCWACAGC	N/A[4]	460	95C for 5 m 95C for 1 m x 35 49C for 1 m x 35 72C for 3 m x 35 72C for 10 m	Luton *et al.* (2002) Morris *et al.* (2014)
	ML-r/mcrR	TTCATTGCRTAGTTWGGRTA--GTT	N/A[4]			
16S rRNA	515F	GTGNCAGCMGCCGCGGTAA	515-534[5]	450	95C for 10 m 95C for 1 m x 30 55C for 1 m x 30 72C for 1 m x 30 72C for 10 m	Muyzer *et al.* (1993)
	926R	CCGYCAATTYMTTTRAGTTT	926-946[5]			

[1] Positions in the *amoA* gene sequence of *Nitrosomonas europaeae* (McTavish *et al.*, 1993)
[2] Positions in the *nirK* gene sequence of *Alcaligenes faecalis* S-6 (Throbäck *et al.*, 2004)
[3] Positions in the *nirS* gene sequence of *Psuedomonas stutzeri ZoBell* ATCC 14405 (Throbäck *et al.*, 2004)
[4] Positions not described in source
[5] Positions in the 16S rRNA gene sequence of *Escherichia coli* (Brosius *et al.*, 1978)

3.2.4.3 *Ion Torrent Sequencing*

All samples including the procedural negative were prepared for Ion Torrent sequencing from the control core. A multiplexed FASTQ file containing all the sequenced reads was returned from sequencing. See Chapter 2.5.1.1.

3.2.5 16S rRNA Gene Sequencing Analysis

Pipeline analysis of sequencing data generated by both Ion Torrent and Illumina 16S rRNA gene sequencing was performed in QIIME 2 (version 2018.6) (Caporaso *et al.*, 2010), using a bespoke set of scripts, which can be found at https://github.com/peterleary/q2pipeline. A full description of the pipeline is detailed in the Methods Chapter and on the GitHub page. Subsequent generation of figures and statistics was performed in R 3.5.1.

3.2.6 Tax4Fun Analysis

The closed-reference OTU Table produced by the pipeline using Vsearch was used as inputs into the Tax4Fun package in R (Asshauer *et al.*, 2015) to create a profile of inferred abundances of KEGG Ortholog genes based on OTU taxonomy. A predefined set of genes of interest were targeted in the Tax4Fun profile (Table 2.3). Profiles of these genes as a function of depth were produced in R 3.5.1. AOV and Tukey HSD was performed in R 3.5.1 on the inferred fractional abundances of the target genes, to identify statistical differences in inferred gene abundances across and between sample depth groups and pairs of depth groups. The four genes targeted by 'end point' PCR and the corresponding inferred abundances of those genes from Tax4Fun were compared by way of XY scatterplot and Spearman correlation, to determine the strength of the association between abundances of the same gene between the two methods by ρ and p-values in R 3.5.1. See Chapter 2.6.

3.2.7 Shotgun Metagenome Sequencing

Six samples from the contaminated core were selected based on previously determined and analysed 16S rRNA gene community patterns, and diluted template DNA extractions were submitted to Northumbria University for sequencing on the Illumina MiSeq platform, using the NextEra XT DNA Library Preparation Kit as per manufacturer's instructions and sequenced using the MiSeq V3 600 bp reagent cartridge (Illumina, USA). Demultiplexed FASTQ files for each sample were returned from Northumbria University. See Chapter 2.5.1.3.

3.2.8 GraftM Analysis

GraftM (Body *et al.*, 2018) was performed on the raw shotgun metagenome sequencing data to identify the presence, abundance, and phylogeny of a set of target genes (Table 3.2) using both precompiled packages as well as a package compiled for the *dsrA* gene.

Table 3.2. A list of functional genes that were targeted using precompiled reference packages in GraftM.

Functional Gene Target	KEGG KO
16S (Greengenes 13.8)	K01977
Nitrate reductase – *narG*	K00370
NiFe hydrogenase (cytochrome c3 hydrogenase) – *hydB*	K00437
Acetyl CoA synthase – *acsB*	K14138
Adenylylsulfate reductase – *aprA*	K00394
Methyl Coenzyme M reductase – *mcrA*	K00399

3.2.9 Geochemical Analysis

The contaminated core was prepared for solid phase elemental analysis by X-ray fluorescence (XRF). See Chapter 2.9.

3.2.9.1 *X-Ray Absorption Spectroscopy*

Selected samples (depths 1, 5, 10, 15, 20, 23, 25, 27, 30 and 35 cm) were analysed by U L3-edge (17166 eV) XAS to determine variation in U oxidation state and coordination environment with depth. See Chapter 2.9.1.

3.2.10 Environmental Fitting

Detrended correspondence analysis (DCA) were produced as per 3.2.5, with geochemical data correlations imposed using ‘envfit’ from the ‘vegan’ package in R 3.5.4. See Chapter 2.8.2.

3.3 Results

3.3.1 DNA Extractable-Quantities and Total Organic Carbon as a Function of Depth

The quantities of DNA extracted from individual depth increments for both cores generally decreased with increasing depth. The amount of DNA extracted from the contaminated core ranged from 46.3 ng/μL at 1 cm to 7.6 ng/μL at 41 cm, with a peak of 57 ng/μL at 13 cm and a low of 3.68 ng/μL at 39 cm (Figure 3.1A). One-way analysis of variance (AOV) on DNA quantities across the four sample depth groups showed a statistical difference (*p* value > 0.05) across all groups, while *post hoc* Tukey HSD between pairs of sample depth groups showed statistical differences in quantities between sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.1A).

DNA quantities from the control core were generally lower than observed in the contaminated core and decreased more rapidly with depth. Values ranged from a peak of 26.4 ng/μL at 1 cm to 3.56 ng/μL at 38 cm (Figure 3.3). AOV showed a statistical difference in DNA quantities across all sample depth groups, and Tukey HSD showed a statistical difference between all pairs of sample depth groups, excluding group pair 3 – 4 (Appendix A.2).

Total organic carbon (TOC) was measured at several depth-increments from the contaminated core, and decreased with depth similar to DNA quantities, with a resurgence at depth. TOC volumes ranged from 84.3% at 5 cm to 13.5% at 34 cm and increased to 36.6% at 40 cm (Figure 3.1B). AOV showed a statistical difference in TOC values across all sample depth groups, and Tukey HSD showed a statistical difference between all pairs of sample depth groups excluding pair 3 – 4 (Appendix A.1B). Spearman correlation of DNA quantities and TOC volumes for each depth sample from the contaminated core showed a statistical positive correlation between DNA and TOC values (Figure 3.2).

3.3.2 Qualitative Assessment of Functional Gene Distributions as a Function of Depth by 'End Point' PCR

'End point' PCR band intensities were quantified via BioNumerics. Band intensities are presented as a qualitative inference of gene abundance within each depth sample, *i.e.*, a higher band intensity denotes a higher gene abundance within a specific sample.

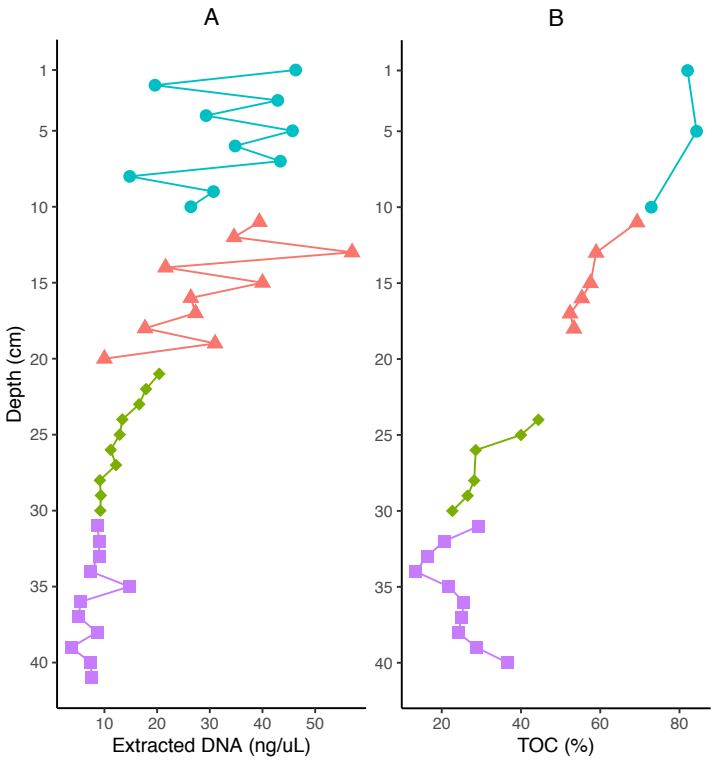


Figure 3.1. Contaminated Core: Depth profiles showing: (A) The amount of DNA extracted from each depth sample from the core in ng/μL, and; (B) The total organic carbon (TOC) measured by Loss on Ignition (LOI) for the depth samples measured. Point colours and shapes denote sample depth group: Blue circles = Group 1 (0 – 10 cm); Red triangles = Group 2 (11 – 20 cm); Green diamonds = Group 3 (21 – 29 cm), and; Purple squares = Group 4 (31 – 41 cm).

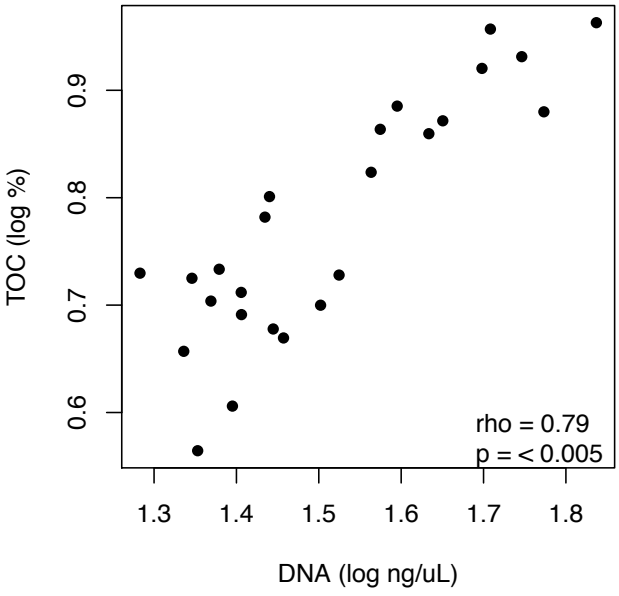


Figure 3.2. Contaminated Depth Core: XY scatterplot showing the log-transformed amount of DNA extracted from each sample against the log-transformed TOC value for the same sample. Spearman correlation is presented as ρ and *p* values.

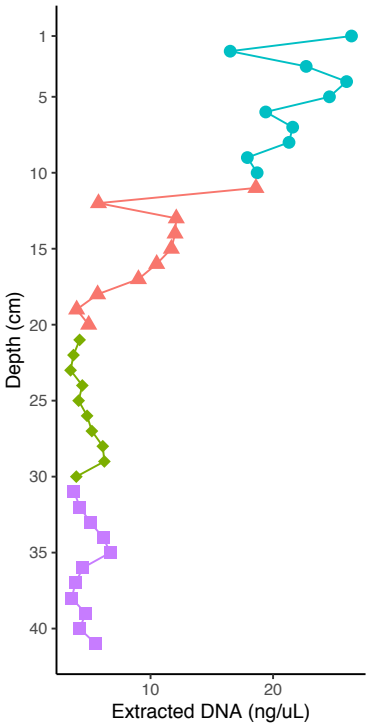


Figure 3.3. Control Core: Depth profiles showing the amount of DNA extracted from each depth sample from the core in ng/μL. Point colours and shapes denote sample depth group: Blue circles = Group 1 (0 – 10 cm); Red triangles = Group 2 (11 – 20 cm); Green diamonds = Group 3 (21 – 29 cm), and; Purple squares = Group 4 (31 – 41 cm).

3.3.2.1 *Contaminated Core*

The bacterial ammonia monooxygenase (*amoA*) gene had relatively low band intensities (compared to other genes targeted) and appeared to be confined to the upper 1 – 13 cm, ranging from a band intensity of 131 at 1 cm to 77 at 13 cm, and was zero between 14 – 37 cm, before re-emerging at 38 – 41 cm with an average intensity of 62.7 ± 0.9 (Figure 3.4A). AOV showed a statistical difference in bacterial *amoA* across all sample depth groups, and Tukey HSD showed a statistical difference between sample depth group pairs 1 – 2, 1 – 3, and 1 – 4 (Appendix A.3A). The archaeal *amoA* gene had higher band intensities than the bacterial counterpart. Band intensities initially increased from a band intensity of 82 at 1 cm to a peak of 254 at 7 cm, and then decreased with depth to 30 at 41 cm (with a low of 29 at 36 cm) (Figure 3.4B). AOV showed a statistical difference in archaeal *amoA* band intensities across all sample depth groups while Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.3B). The nitrite reductase *nirS* appeared to persist with depth in the core from surface to ~ 26 cm, and sporadic intervals thereafter. The average band intensity for *nirS* between 1

– 26 cm was 228.3 ± 2.3, at 27 cm, band intensity lowered to 155, and then to an average of 47 ± 5.9 between 28 – 34 cm. Band intensities ranged from 234 at 35 cm to 35 at 36 cm and averaged 105.5 ± 17.3 between 37 – 41 cm (Figure 3.4C). AOV showed a statistical difference across all sample depth groups, and Tukey HSD showed a statistical difference between sample depth group pairs 1 – 4, 2 – 4, and 3 – 4 (Appendix A.3C). The other nitrite reductase, *nirK*, decreased more progressively with depth than the *nirS*, ranging from a peak band intensity of 246 at 4 cm and declined gradually to a low of 61 at 33 cm. Similar to the *nirS*, the *nirK* peaked again at 35 cm to a band intensity of 237 and persisted between 36 – 41 cm (average band intensity 174 ± 21.4) (Figure 3.4D). AOV showed a statistical difference in *nirK* band intensities across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3 and 1 – 4 (Appendix A.3D). The dissimilatory sulfate reductase *dsrA* gene was split between two halves of the core, having consistently lower band intensities in the upper 1 – 18 cm than the lower 19 – 41 cm. Band intensity averaged 125.9 ± 5.5 between 1 – 18 cm, and averaged 220 ± 6.7 between 19 – 41 cm, including a sudden drop to 136 at 34 cm (Figure 3.4E). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey HSD showed statistical differences between sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.3E). The methyl coenzyme M reductase *mcrA* gene was present in the 1 – 7 cm region, with an average band intensity of 105 ± 11.6, which decreased to 56 ± 5.2 between 8 – 21 cm, between 22 – 26 cm increased to 173 ± 16 and fluctuated between 27 – 41 cm from 80 – 170 with an average band intensity of 129 ± 6.9 (Figure 3.4F). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey HSD showed statistical differences between sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.3F).

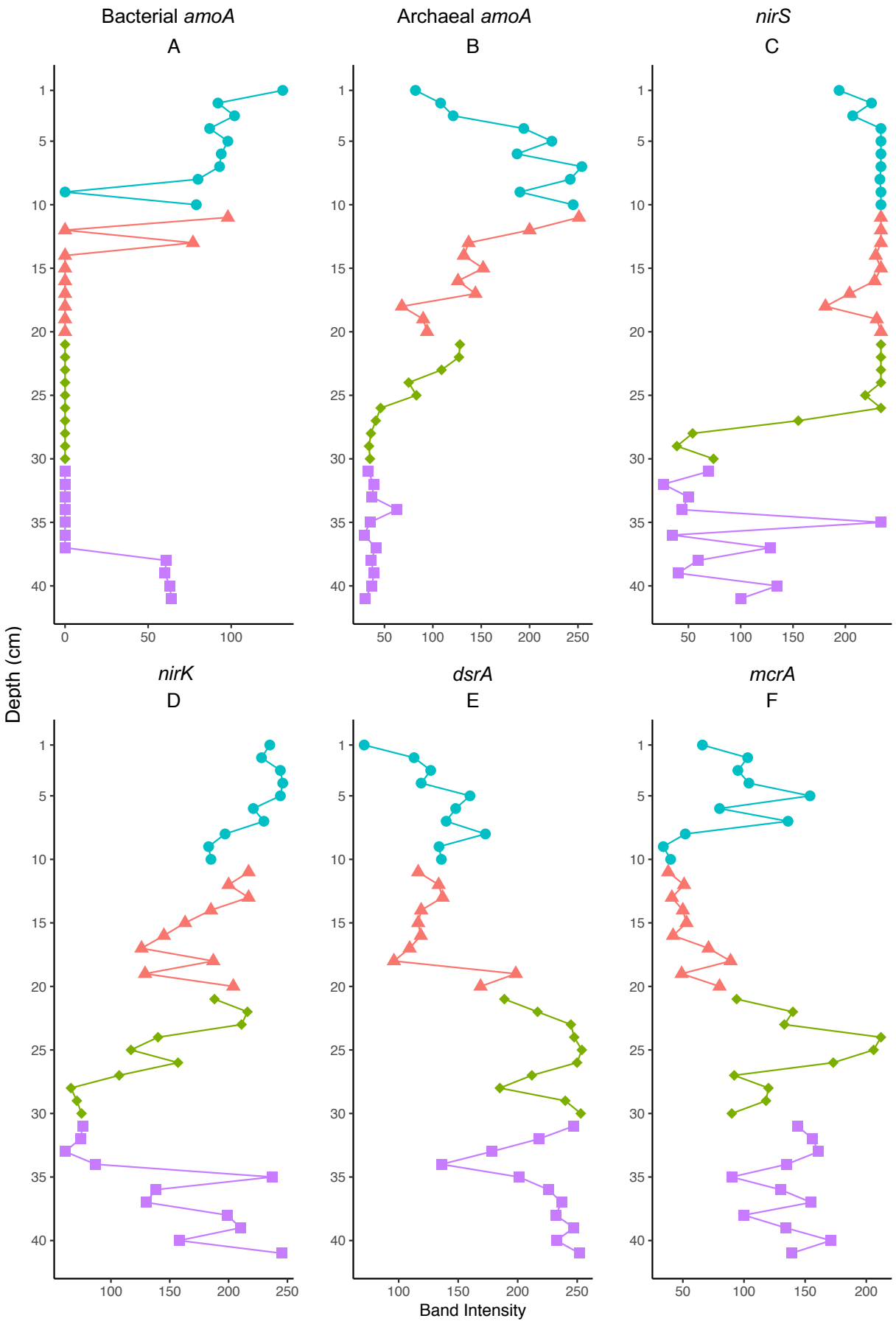


Figure 3.4. Contaminated Core: Depth profiles of 'end point' PCR target gene band intensities, as determined by agarose gel electrophoresis and BioNumerics, as a function of depth in cm. Colour shapes refer to sample depth group. (A). Bacterial *amoA*; (B). Archaeal *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA*.

3.3.2.2 *Control Core*

The bacterial *amoA* gene had similarly low band intensities as was observed with the contaminated core, and generally decreased with depth albeit did not decrease to zero as in the contaminated core. Band intensity peaked at 45 at 1 cm and decreased to a low of 23 at 36 cm (Figure 3.5A). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey HSD showed statistical differences between sample depth group pairs 1 and all other groups (Appendix A.4A). The archaeal *amoA* generally had higher band intensities than the bacterial *amoA*, and decreased progressively with depth, as was observed in the contaminated core. Band intensities ranged from a peak of 193 at 5 cm to 33 at 37 cm (Figure 3.5B). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey HSD showed statistical differences between sample depth group pairs 1 and all other groups (Appendix A.4B). The *nirS* gene was confined to the upper 1 – 12 cm region, much shallower than the contaminated core. The average band intensity between 1 – 12 cm was 162 ± 9, which lowered to 43 ± 2.8 between 13 – 41 cm (Figure 3.5C). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey HSD showed statistical differences between sample depth group pairs 1 and all other groups, and pair 2 – 4 (Appendix A.4C). The *nirK* band intensities were very low in the control core, but generally decreased with depth from a peak band intensity of 86 at 4 cm to a low of 33 at 37 cm (Figure 3.5D). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey HSD showed statistical differences between sample depth group pairs 4 and all other groups (Appendix A.4D). The *dsrA* gene increased in band intensity more gradually in the control core than in the contaminated core, increasing from a band intensity of 48 at 1 cm to 86 at 8 cm, and generally persisted throughout the remainder of the core (average band intensity of 119 ± 5.1) (Figure 3.5E). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey HSD showed statistical differences between sample depth group pairs 1 – 2, 1 – 3, and 3 – 4 (Appendix A.4E). Similar to the *dsrA* gene, band intensities of the *mcrA* gene increased progressively from 1 – 18 cm from a band intensity of 49 at 1 cm to 146 at 18 cm. This then decreased to an average band intensity of 76 ± 3.7 between 19 – 25 cm, and from 26 – 41 cm, band intensity fluctuated from 56 at 41 cm to 138 at 35 cm (Figure 3.5F). AOV showed a statistical difference in band intensities across all sample depth groups, and Tukey

HSD showed statistical differences between sample depth group pairs 1 – 2, 2 – 3, and 2 – 4 (Appendix A.4F).

3.3.3 16S rRNA Gene Sequencing Analysis Results

16S rRNA gene sequencing produced 1,875,571 paired-end reads of 251 nt for the contaminated core and 2,835,681 single-end reads with a median 409 nt for the control core. Pipeline analysis using QIIME2 produced 18,232 amplicon sequence variants (ASVs) from 1,545,680 reads of 227 nt for the contaminated core, and 11,967 ASVs from 2,290,821 reads of 230 nt for the control core.

3.3.3.1 *Alpha Diversity Analysis*

Shannon's species richness diversity appeared to decrease with depth in both the contaminated (Figure 3.6A) and control (Figure 3.7A) depth cores. AOV showed a statistical difference in Shannon diversity indices across the sample depth groups of both cores, while Tukey HSD showed statistical differences between sample depth group pairs 1 and all other groups, pair 2 – 4, and in the control core pair 3 – 4 (Appendix A.5A and 9B). Faith's phylogenetic diversity did not appear to change as a function of depth in either core (Figure 3.6B and 7B), despite borderline-significant *p* values from AOV of 0.063 and 0.071 for the contaminated and control cores respectively, and Tukey showing a statistical difference in indices values between sample depth group pair 3 – 4 (Appendix A.5C and 9D). Pielou's species evenness decreased with depth in both cores (Figure 3.6C and 7C), with AOV resulting in a statistical difference in index values across all sample depth groups in both cores, and Tukey showing statistical differences between sample group pairs 1 and all other groups, 2 – 4, and 3 – 4 (Appendix A.5E and 9F).

3.3.3.2 *Beta diversity*

Non-metric multidimensional scaling (nMDS) based on Bray Curtis dissimilarity (Figure 3.8) and detrended correspondence analysis (DCA) plots (Figure 3.9) based on amplicon sequence variant (ASV) assemblage of the samples from both cores appeared to show distinct separation of the two cores from one another,

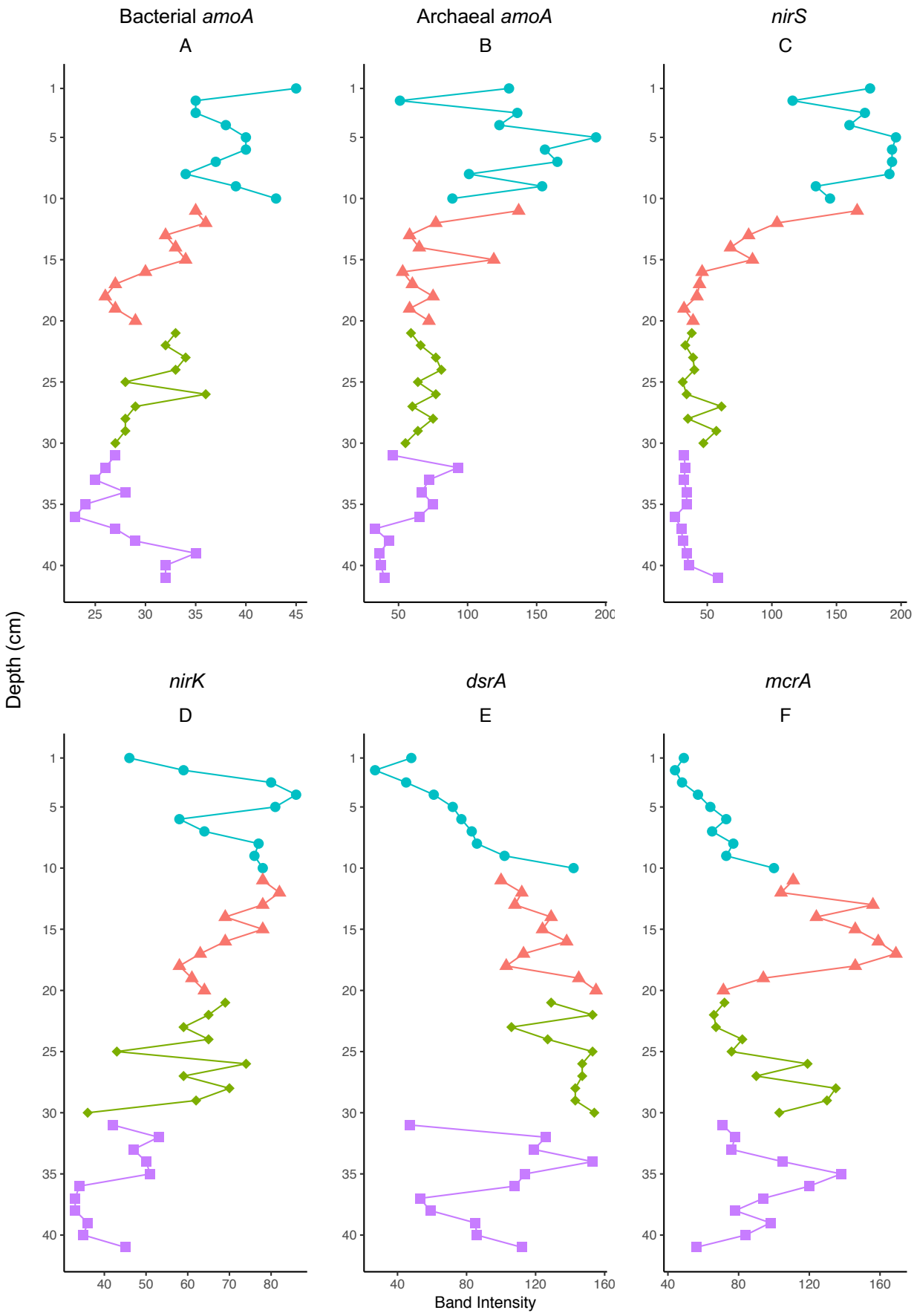


Figure 3.5. Control Core: Depth profiles of 'end point' PCR target gene band intensities, as determined by agarose gel electrophoresis and BioNumerics, as a function of depth in cm. Colour shapes refer to sample depth group. (A). Bacterial *amoA*; (B). Archaeal *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA*.

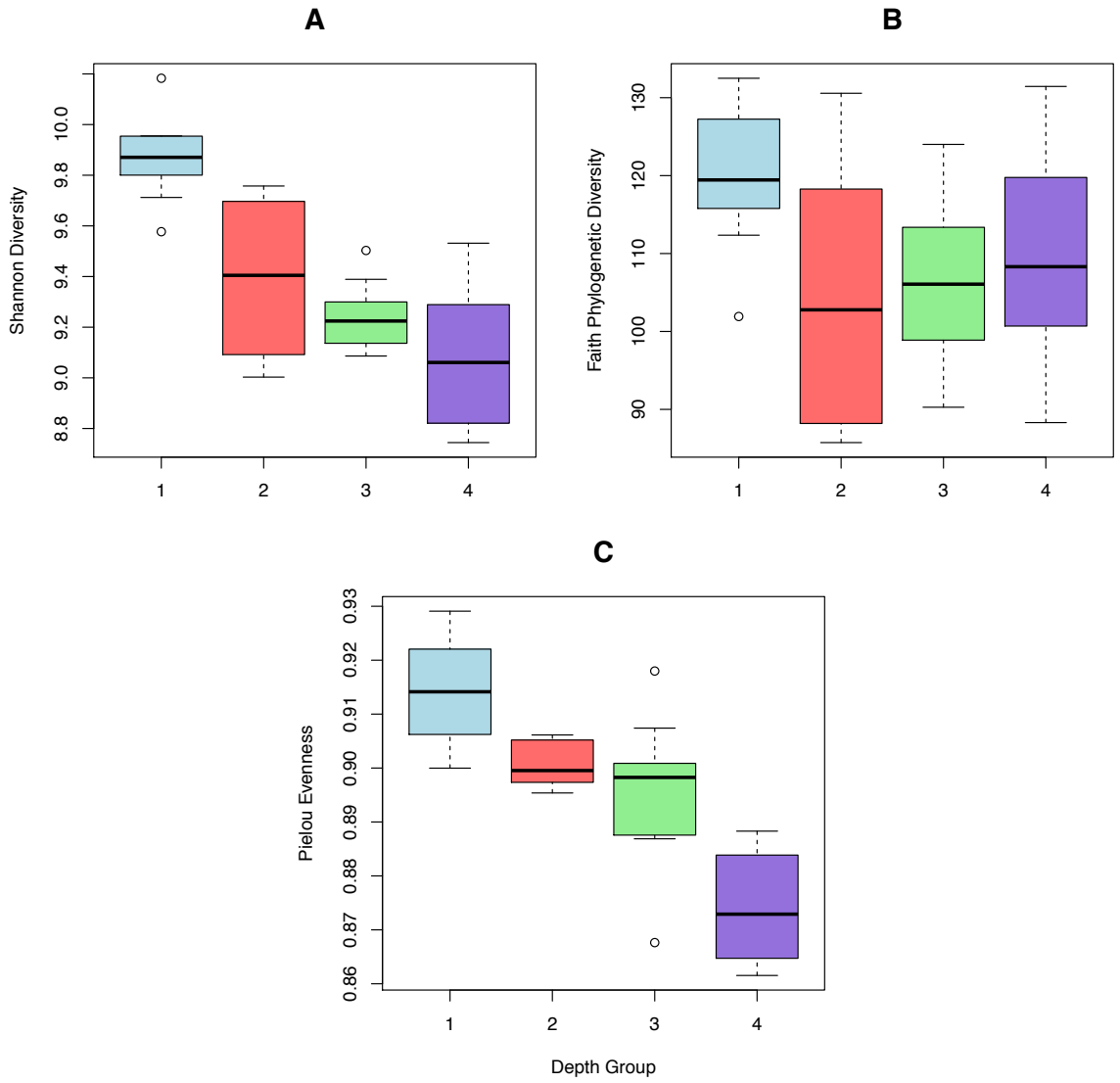


Figure 3.6. Contaminated Core: Whisker boxplots denoting range in alpha diversity indices values of the four depth sample groups from the Needle's Eye contaminated core. Colours denote depth group sample: Blue = Group 1 (0 – 10 cm); Red = Group 2 (11 – 20 cm); Green = Group 3 (21 – 29 cm), and; Purple = Group 4 (31 – 41 cm). (A). Shannon Species Richness; (B) Faith Phylogenetic Diversity; (C) Pielou Species Evenness.

but samples from both cores distributed in order of increasing depth in the same order in the same fashion along the nMDS and DCA axes. Both plots appear to show samples 1 ~ 20 cm forming a line on the plots, while samples from ~ 22 – 41 cm cluster together more closely. This is especially true of the contaminated core samples on the DCA plot (Figure 3.9).

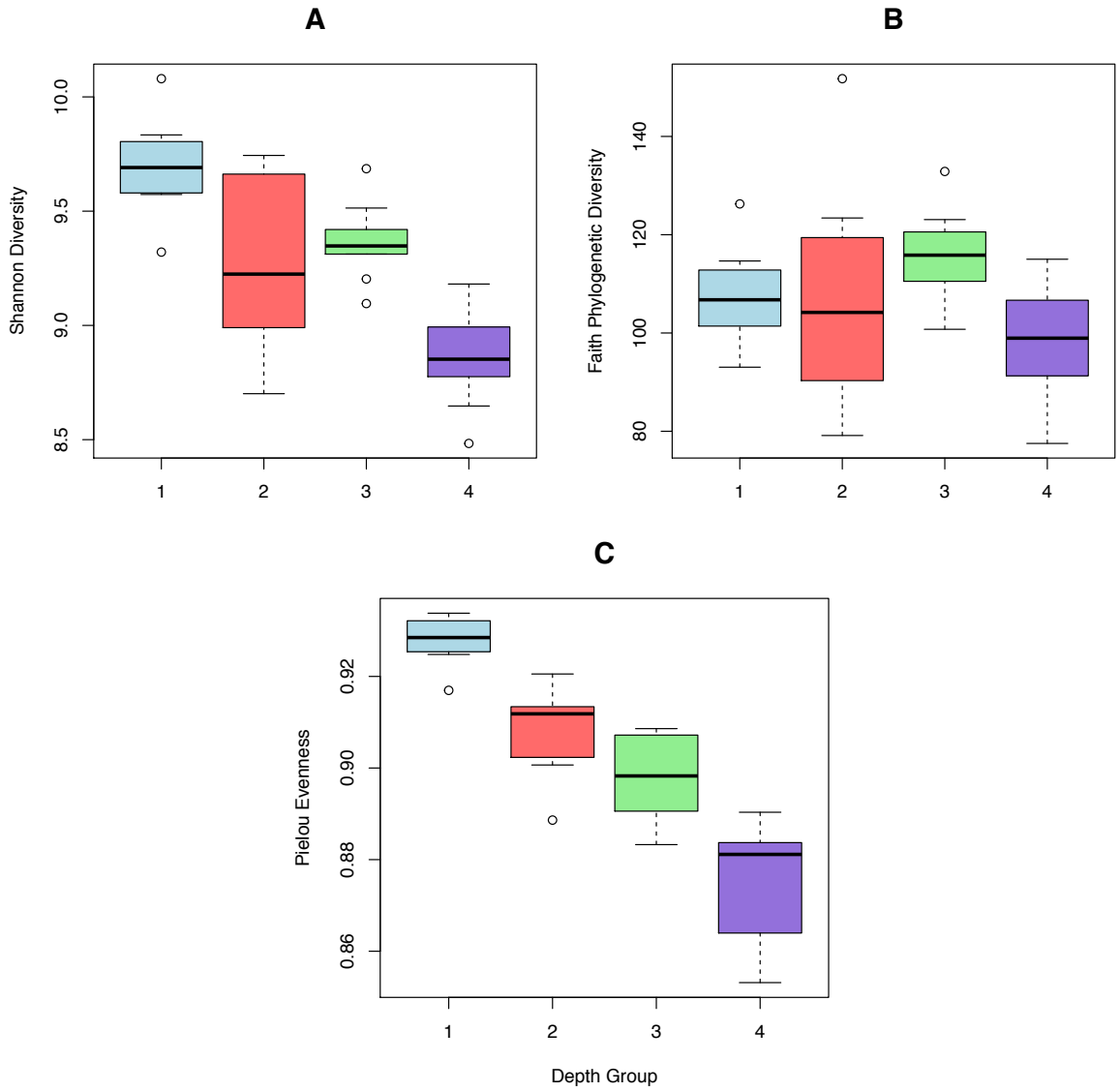


Figure 3.7. Control Core: Whisker boxplots denoting range in alpha diversity indices values of the four depth sample groups from the Needle's Eye contaminated core. Colours denote depth group sample: Blue = Group 1 (0 – 10 cm); Red = Group 2 (11 – 20 cm); Green = Group 3 (21 – 29 cm), and; Purple = Group 4 (31 – 41 cm). (A). Shannon Species Richness; (B) Faith Phylogenetic Diversity; (C) Pielou Species Evenness.

The nMDS (Figure 3.10) and DCA (Figure 3.11) plots of the contaminated core community assemblages appear to show distinct separation of the two halves of the depth core, comprised of 1 – 23 cm and 24 – 41 cm separately. 35 cm appears to be acting as an outlier from either group. Both plots appear to show linear depth-related distribution of samples 1 – 23 cm along the plot axes, while samples 24 – 41 appear to form a less depth-defined cluster of samples. 35 cm appears to be an outlier from both clusters of samples.

The nMDS (Figure 3.12) and DCA (Figure 3.13) plots of the control core community assemblages appear to show similar depth-related distribution of the samples albeit with less of a defined cut-off between two halves of the core as observed in the contaminated core. Here, distribution of samples is more progressive with depth, and the nMDS appears to exhibit the arching effect commonly observed in large, sparse microbial datasets that can mask true community assemblage diversity between highly-differentiated samples (Morton *et al.*, 2017). This arching is not observed in the DCA. In the control core nMDS and DCA plots, similar to the control core, the upper 1 ~ 20 cm samples follow a depth-related distribution that is spread out across the axes, while the 21 – 41 cm samples form more of a cluster to one side of the plots.

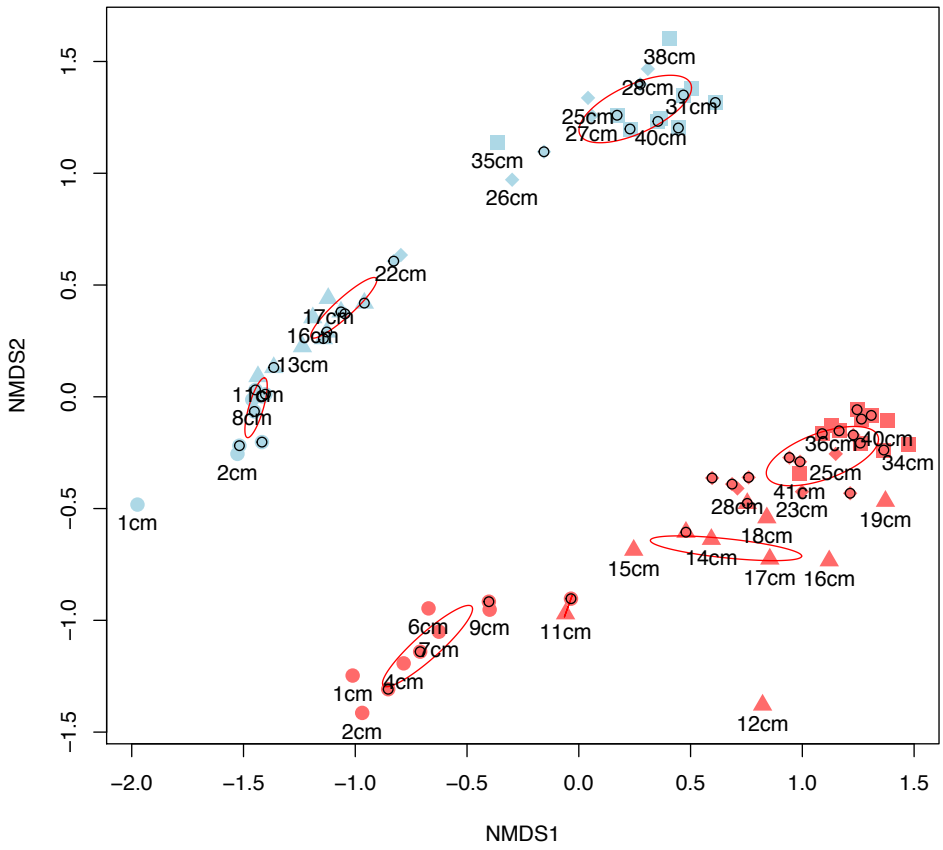


Figure 3.8. Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity metric of samples from both depth cores. Blue denotes contaminated core samples and red denotes control core samples. Point shapes denote sampling depth groups. Black circles denote samples without a label (due to space constraints). Red outline circles denote hierarchical clustering of samples based on centroid similarity.

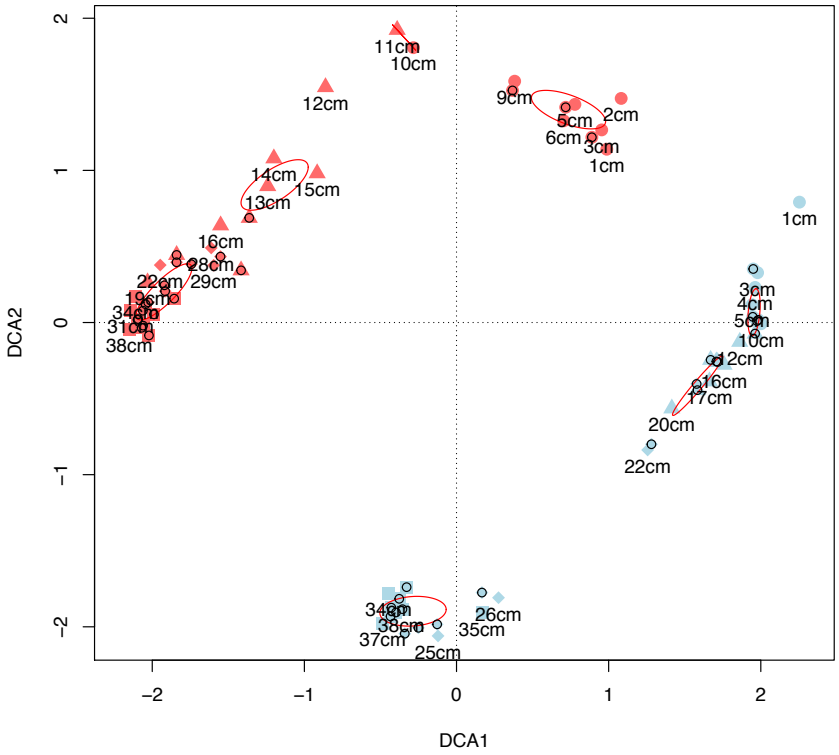


Figure 3.9. Detrended correspondence analysis (DCA) of samples from both depth cores. Blue denotes contaminated core samples and red denotes control core samples. Red outline circles denote hierarchical clustering of samples based on centroid similarity.

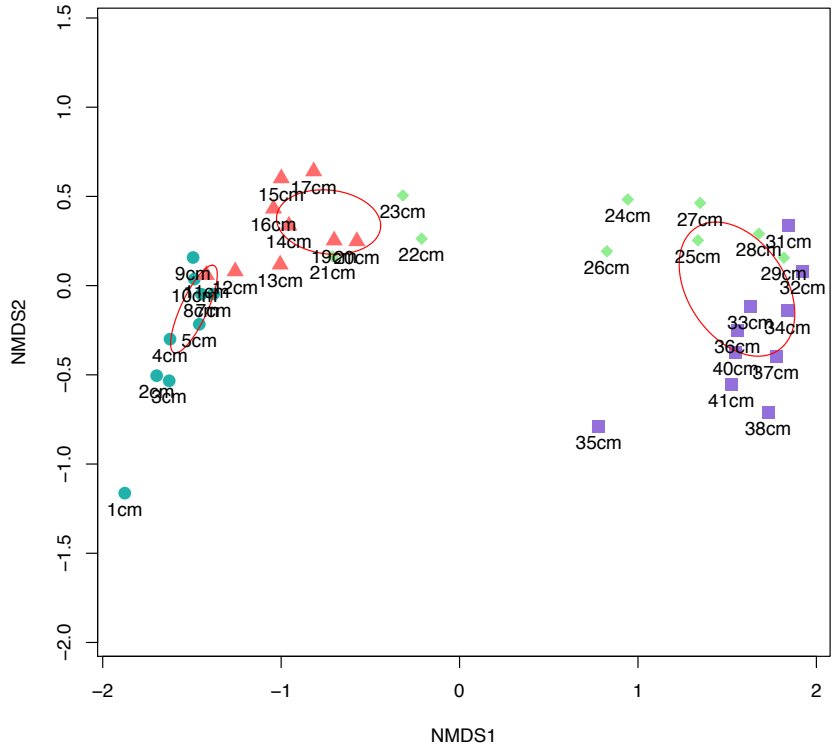


Figure 3.10. Contaminated Core: nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Colours and points denote sample depth groups.

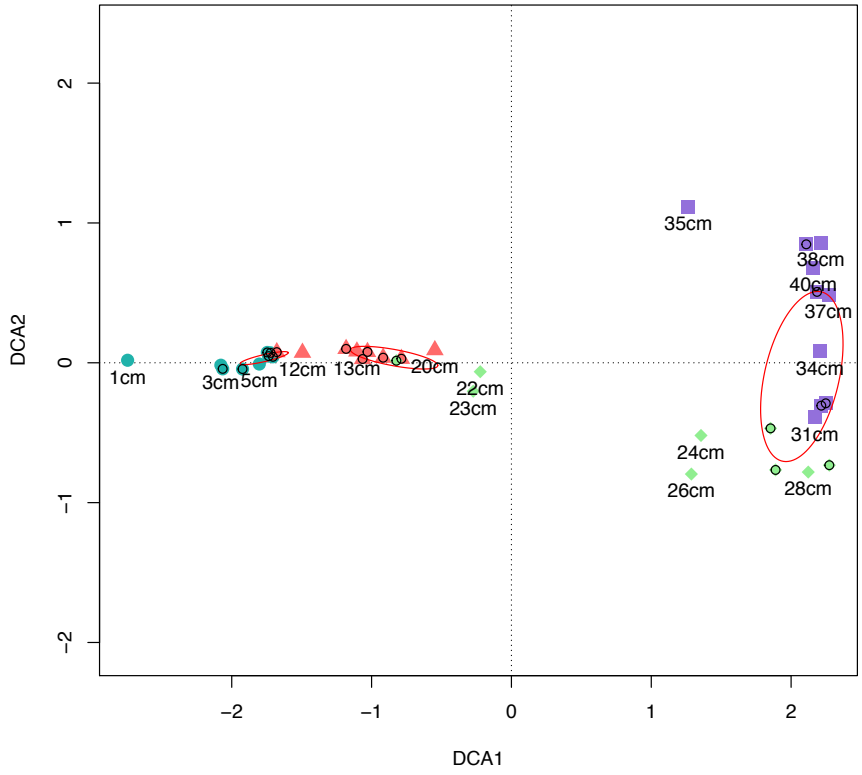


Figure 3.11. Contaminated Core: DCA plot of microbial community assemblages of samples. Colours and points denote sample depth groups.

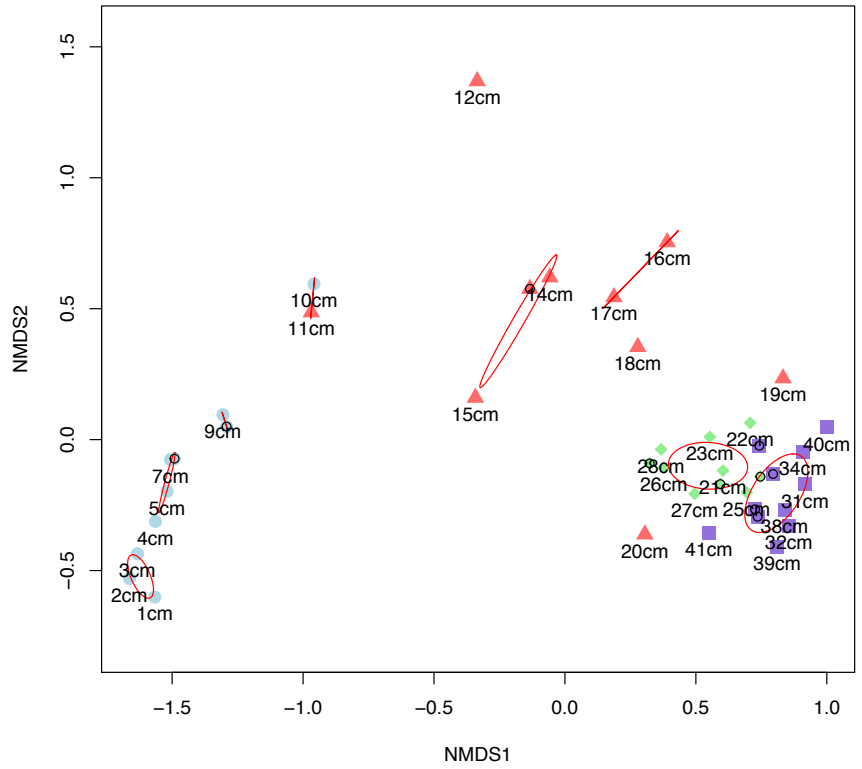


Figure 3.12. Control Core: nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Colours and points denote sample depth groups.

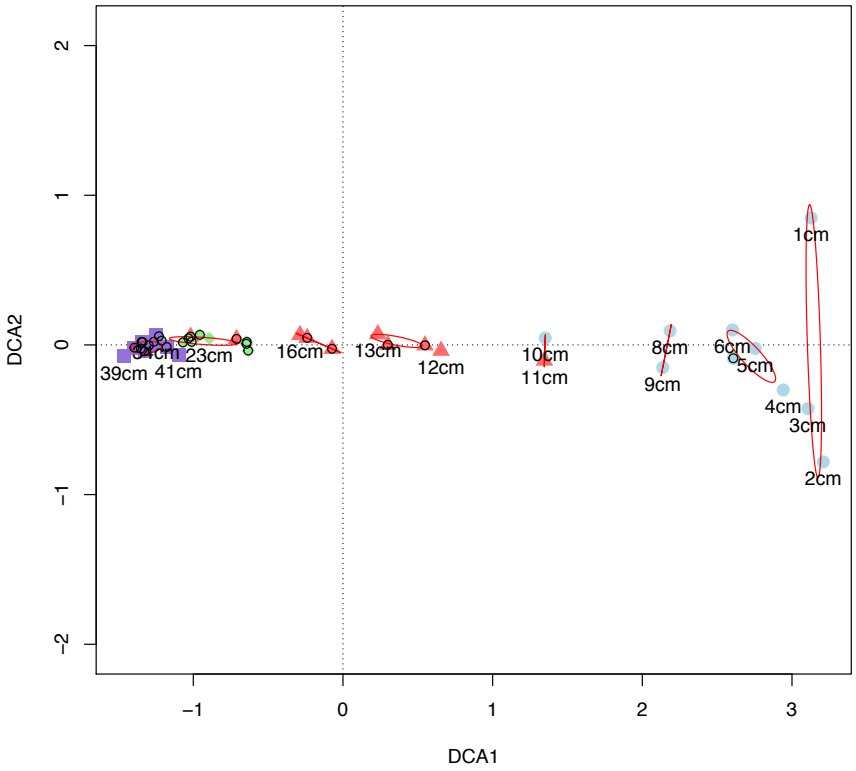


Figure 3.13. Control Core: DCA plot of microbial community assemblages of samples. Colours and points denote sample depth groups.

3.3.4 Taxonomic Analysis

3.3.4.1 *Contaminated Core Taxonomic Profiles*

The taxonomic profile of depth samples from the contaminated core produced by 16S rRNA gene sequencing pipeline analysis in QIIME2 appeared to show diverse microbial communities comprised of many taxa. There were 1,651 genera. The ten most abundant ASVs collapsed at genus level are presented in Figure 3.14, and these ten genera represented between 12.6% (1 cm) and 30% (19 cm) of the respective microbial communities at genus level. The most abundant genus was an uncultured *Acidobacteria* subgroup 6, which was most abundant in the upper region of the core and decreased with depth, peaking in abundance at 5 cm with 9.39% relative frequency, and a low of 0.32% at 37 cm, with a peak of 2.18% at 35 cm. AOV showed a statistical difference in *Acidobacteria* subgroup 6 genus abundance across the four sample depth groups, and Tukey HSD showed statistical differences between all sample depth group pairs excluding pair 1 – 2 (Appendix A.6A). The second most abundant genus in the core was an uncultured *Anaerolineaceae*. This appeared to increase with depth from 1 – 27 cm (0.97% to 6.45% at 1 and 27 cm

respectively), before decreasing from 28 – 41 cm (5.04% at 28 cm to 3.81% at 41 cm), including a dip at 35 cm to 2.92%. AOV showed a statistical difference in *Anaerolineaceae* ASV abundance across all sample groups, and Tukey HSD showed statistical differences between all group pairs except 1 – 4 and 2 – 3 (Appendix A.6B). Next, an uncultured 'methanogenic' *Bathyarchaeia* genus appeared to increase in abundance with depth, from 0.39% at 1 cm to 4.4% at 24 cm and decreased from 4.17% at 25 cm to 3.49% at 41 cm, including a dip at 35 cm to 1.97%. AOV showed a statistical difference in 'methanogenic' *Bathyarchaeia* abundance across all sample groups, and Tukey HSD showed statistical differences between all group pairs except pair 3 – 4 (Appendix A.6E). An uncultured *Chloroflexi* KD4-96 appeared to decrease with depth, ranging from 3.27% at 1 cm to 3.9% at 19 cm, before decreasing from 3.64% at 20 cm to 0.9% at 34 cm, rising to 2.45% at 35 cm, and decreasing to 1.42 ± 0.22% between 36 – 41 cm. AOV showed a statistical difference in *Chloroflexi* KD4-96 abundance across all sample group and Tukey HSD showed statistical differences in abundance between sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.6C). The genus *Sulfurifustis* increased in abundance from 0.24% at 1 cm to 4.55% at 20 cm and decreased from 4.01% at 21 cm to 1.06% at 41 cm, with a spike at 35 cm to 3.16%. AOV showed a statistical difference in *Sulfurifustis* abundance across all sample depth groups and Tukey HSD showed statistical differences between sample group pairs 1 – 2, 2 – 4, and 3 – 4 (Appendix A.6D). An uncultured δ-*proteobacteria* Sva0485 genus appeared to increase with depth, with an average abundance of 0.19 ± 0.06% between 1 – 23 cm and included zero counts between 1 – 7, 9, and 11 cm, and increased to 3.68% at 24 cm, and 4.49 ± 0.38% between 25 – 41 cm. AOV showed a statistical difference in δ-*proteobacteria* Sva0485 abundance across all sample groups, and Tukey HSD showed statistical differences between all pairs of sample groups excluding pair 1 – 2 (Appendix A.6F). An uncultured NC10 *Rokubacteriales* appeared to increase from 1.24% at 1 cm to 4.26% at 8 cm, and decrease from 3.31% at 9 cm to 0% at 25 cm, and from 26 – 41 cm, had an average abundance of 0.57 ± 0.12%, including a spike to 1.43% at 35 cm. AOV showed a statistical difference in NC10 *Rokubacteriales* abundance across all sample groups, and Tukey HSD showed statistical differences between all sample group pairs excluding 1 – 2 and 3 – 4 (Appendix A.6G). Another uncultured δ-*proteobacteria*, the MBNT15, increased in abundance from 0.37% at 1 cm to 3.87% at 24 cm, and generally declined slightly from 2.82% at 25 cm to 1.24% at 41 cm. AOV showed a statistical difference in δ-*proteobacteria* MBNT15

abundance across all sample depth groups, and Tukey HSD showed a statistical difference between sample group pairs 1 and all other groups, and 3 – 4 (Appendix A.6H). An uncultured

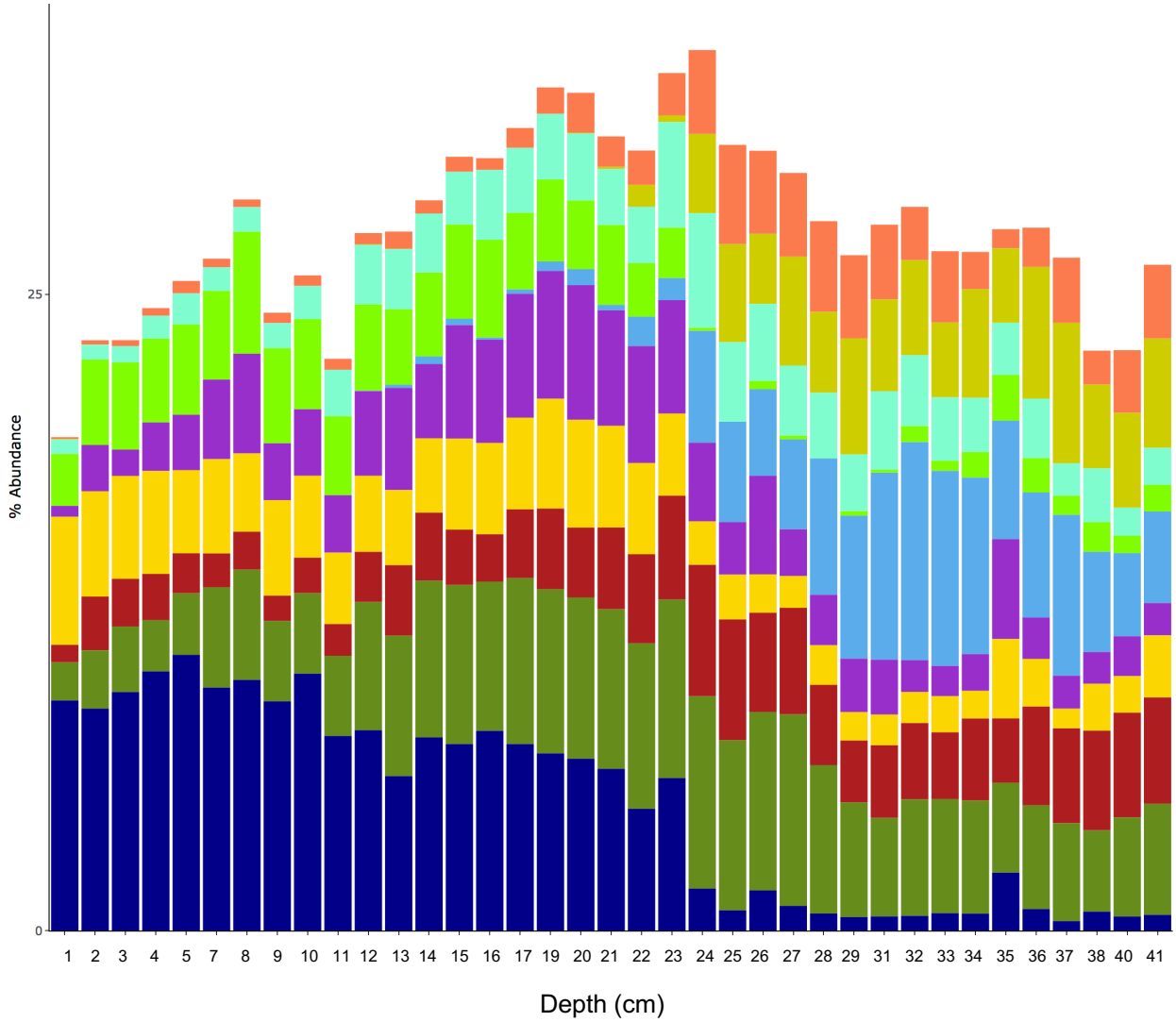


Figure 3.14. Contaminated Core: 100% stacked bar chart of the ten most abundant ASV taxonomies collapsed at genus level. Size of bar denotes percentage abundance of the genera as a proportion of total ASV genera assemblage within each sample. Dark blue = uncultured *Acidobacteria* subgroup 6; dark green = uncultured *Anaerolineaceae*; red = uncultured methanogenic *Bathyarchaeia*; yellow = uncultured *Chloroflexi* KD4-96; purple = *Sulfurifustis* spp.; blue = uncultured δ-*proteobacteria* Sva0485; light green = uncultured *Rokubacteriales* NC10; light blue = uncultured δ-*proteobacteria* MBNT15; gold = uncultured *Thermoprotei* archaeon, *Bathyarchaeia*; orange = uncultured *Thermodesulfovibriona*.

Thermoprotei ASV from the class *Bathyarchaeia* was apparently confined to the lower region of the core, where between 1 – 23 cm the average abundance was 0.05 ± 0.04%, including zero counts between 1 – 11, and 14 – 19 cm. This increased to 2.6% at 24 cm, and average abundance was 3.28 ± 0.2% between 25 – 41 cm,

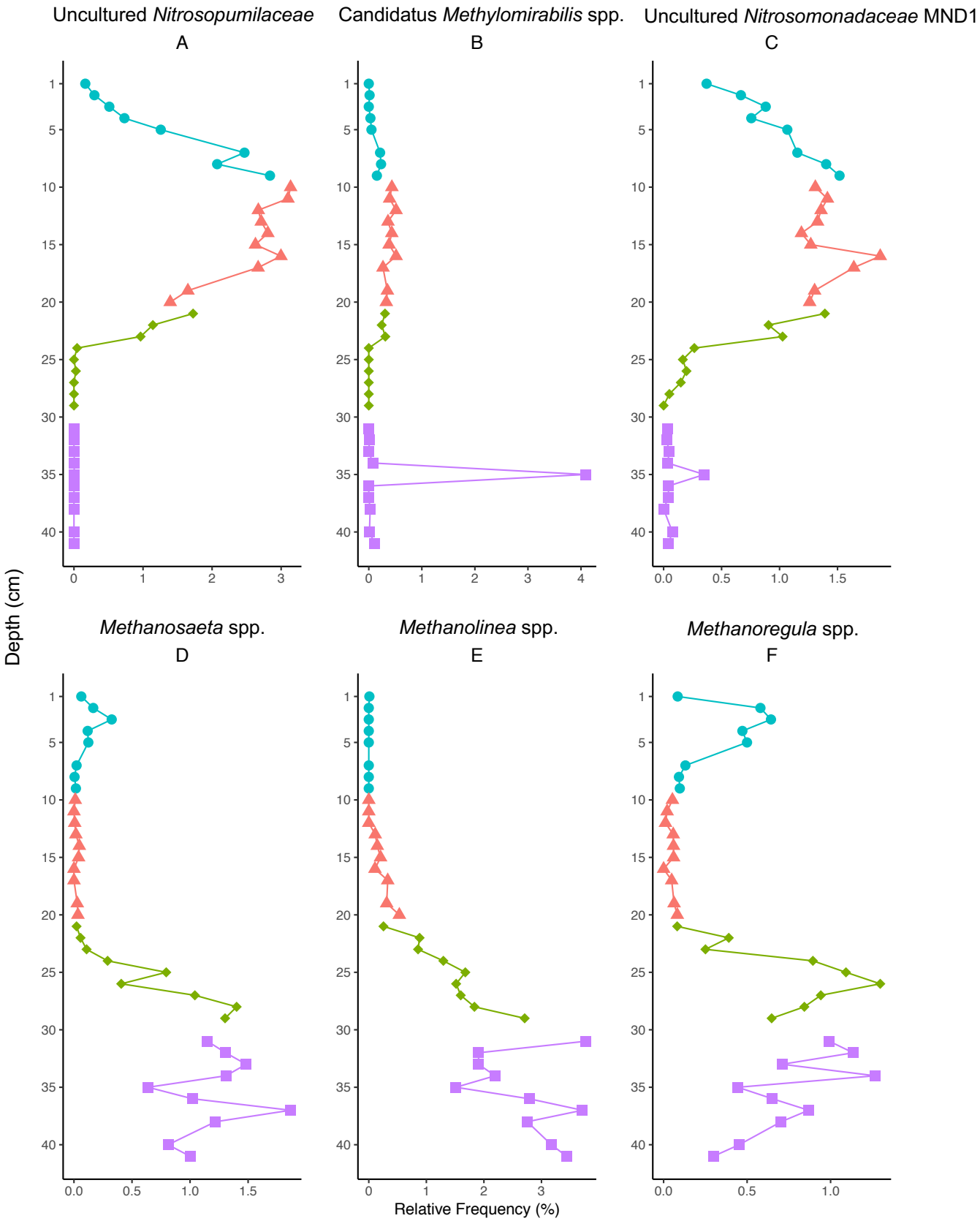


Figure 3.15. Contaminated Core: Depth profiles of six additional key functional microbial taxa. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). Uncultured *Nitrosopumilaceae*; (B). Candidatus *Methylomirabilis* spp.; (C). uncultured *Nitrosomonadaceae* MND1; (D). *Methanosaeta* spp.; (E). *Methanolinea* spp.; (F). *Methanoregula* spp.

including a dip to 2.28% at 35 cm. AOV showed a statistical difference in *Thermoprotei* abundance across all sample depth groups, and Tukey HSD showed a statistical difference between all sample group pairs except 1 – 2 (Appendix A.6I). An uncultured *Thermodesulfovibriona* from the phylum *Nitrospirae* appeared to increase with depth, from 0.04% at 1 cm to 3.27% at 25 cm, and between 26 – 41 cm had an average abundance of 2.06 ± 0.2%, including a dip to 0.58% at 35 cm.

AOV showed a statistical difference in *Thermodesulfovibriona* abundance across all sample groups, and Tukey HSD showed statistical differences between all sample group pairs except 1 – 2 and 3 – 4 (Appendix A.6J).

Six additional ASVs collapsed at genus level were identified as they represented monophyletic taxonomic genera with defined functions of interest. A genus from the *Nitrosopumilaceae* family and the phylum *Thaumarchaeota* increased from 0.16% at 1 cm to 3.14% at 10 cm, decreased from 3.1% at 11 cm to 0.96% at 23 cm, before decreasing further to 0.01 ± 0.003% between 24 – 41 cm, including many zero counts (Figure 3.15A). AOV showed a statistical difference in *Nitrosopumilaceae* abundance across all sample groups, and Tukey HSD showed statistical differences between group pairs 1 – 2, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.6K). The *Candidatus Methylomirabilis* genus was almost exclusively confined to 35 cm (Figure 3.15B), and subsequently AOV and Tukey HSD did not show statistical differences in abundance across or between sample groups (Appendix A.6L). An uncultured *Nitrosomonadaceae* MND1 increased from 0.37% at 1 cm to 1.87% at 16 cm and decreased from 1.64% at 17 cm to 0 at 29 cm, and between 31 – 41 cm, had an average abundance of 0.07 ± 0.03%, including a small spike to 0.35% at 35 cm (Figure 3.15C). AOV showed a statistical difference in *Nitrosomonadaceae* MND1 abundance across all sample depth groups, and Tukey HSD showed statistical differences between sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4, while pairs 1 – 2 and 3 – 4 were both borderline-significant (Appendix A.6M). The *Methanosaeta* spp. had an average abundance of 0.16 ± 0.05% between 1 – 5 cm, which decreased to 0.03 ± 0.01% between 6 – 23 cm and increased from 0.3% at 24 cm to 1.48% at 33 cm, dipped at 35 cm to 0.64%, and peaked at 37 cm to 1.87% (Figure 3.15D). AOV showed a statistical difference in *Methanosaeta* spp. abundance across all sample groups, and Tukey HSD showed a statistical difference between all pairs of sample groups except pair 1 – 2 (Appendix A.6N). The *Methanolinea* spp. had an abundance of 0.01% at 1 cm, and zero counts between 2 – 12 cm, abundance

increased from 0.11% at 13 cm to 3.76% at 31 cm, lowered to 1.51% at 35 cm, and increased again to 3.71% at 37 cm (Figure 3.15E). AOV showed a statistical difference in *Methanolinea* spp. abundance across all sample groups, and Tukey HSD showed a statistical difference between all pairs of sample groups except pair 1 – 2 (Appendix A.6O). The *Methanoregula* spp. had a similar distribution to the *Methanosaeta* spp., albeit with lower abundances. At 1 cm, abundance was 0.08% which increased to 0.64% at 3 cm and decreased to 0.13% at 7 cm. Between 8 – 21 cm, the average abundance was 0.06 ± 0.01%. The abundance increased from 0.39% at 22 cm to 1.3% at 26 cm, after which abundances generally decreased from 0.94% at 27 cm to 0.3% at 41 cm (Figure 3.15F). AOV showed a statistical difference in *Methanoregula* spp. abundance across all sample groups, and Tukey HSD showed a statistical difference between all pairs of sample groups except pair 1 – 2 and 3 – 4 (Appendix A.6P).

3.3.4.2 *Control Core Taxonomic Profiles*

The taxonomic profile of depth samples from the control core produced by 16S rRNA gene sequencing pipeline analysis in QIIME2 appeared to show diverse microbial communities comprised of many taxa. There were 1,347 genera. The ten most abundant ASVs collapsed at genus level are presented in Figure 3.16, and these ten genera represented between 11.26% (2 cm) and 31.8% (34 cm) of the respective microbial communities at genus level. Abundances of many of these taxa changed more progressively as a function of depth compared to the contaminated core, and many of the 10 most abundant taxa in the core exhibited patterns of increased abundance with depth.

Similar to the contaminated core, the most abundant ASV at genus level was the *Acidobacteria* subgroup 6, and also similar to the contaminated core, was most abundant in the upper part of the core. Abundance of the *Acidobacteria* subgroup 6 peaked at 8.39% at 3 cm and decreased gradually to 0.9% at 39 cm. AOV showed a statistical difference in *Acidobacteria* subgroup 6 abundance across all sample depth groups, and Tukey HSD showed statistical differences between sample depth group pairs 1 and all other groups, and 2 – 4 (Appendix A.7A). An uncultured ‘methanogenic’ *Bathyarchaeia* increased gradually in abundance with depth, 0.14% at 2 cm to 5.17% at 38 cm. AOV showed a statistical difference in ‘methanogenic’ *Bathyarchaeia* abundance across all sample depth groups, and Tukey HSD showed statistical differences between sample depth group pairs 1 and all other groups

(Appendix A.7B). An uncultured *Gallionellaceae* increased gradually from 0% at 4 cm to 7.24% at 35 cm. AOV showed a statistical difference in *Gallionellaceae* abundance across all sample groups and Tukey HSD showed statistical differences across all pairs of sample groups (Appendix A.7C). An uncultured α-*proteobacteria* also gradually increased with depth, from 0.15% at 1 cm to 6.47% at 39 cm. A BLAST search of the representative sequence of this ASV showed the best match to be an uncultured *Rhodospirallaceae* clone from a coal seam environment (95% match at 100% coverage). AOV showed a statistical difference in the α-*proteobacteria* abundance across all sample groups and Tukey HSD showed statistical differences between all pairs of sample groups excluding 1 – 2 (Appendix A.7D). An uncultured *Anaerolineaceae* increased from 1.21% at 1 cm to 3.52% at 19 cm, and from 20 – 41 cm had an abundance of 2.83 ± 0.07%. AOV showed a statistical difference in *Anaerolineaceae* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 and all other groups and 2 – 3 (Appendix A.7E). An uncultured *Nitrospirae* 4-29-1 increased progressively with depth from 0.51% at 1 cm to 5.23% at 40 cm. AOV showed a statistical difference in *Nitrospirae* 4-29-1 abundance across all sample groups and Tukey HSD showed statistical differences between sample group pairs 4 and all other groups, and 1 – 3 (Appendix A.7F). An uncultured *Acidobacteriales* had a low abundance between 1 – 9 cm, increasing from 0.16% at 1 cm to 0.76% at 9 cm. This then increased from 1.93% at 10 cm to a peak of 4.43% at 19 cm, and then decreased from 2.86% at 20 cm to 1.49% at 41 cm. AOV showed a statistical difference in *Acidobacteriales* abundance across all sample groups and Tukey HSD showed statistical differences between all sample group pairs (Appendix A.7G). An uncultured δ-*proteobacteria* MBNT15 increased from 0.3% at 1 cm to 0.8% at 8 cm, and from 1.72% at 9 cm to 4% at 15 cm, and then decreased from 2.68% at 16 cm to 1.65% at 41 cm. AOV showed a statistical difference in δ-*proteobacteria* MBNT15 abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2, 1 – 3, 2 – 3, and 2 – 4 (Appendix A.7H). An uncultured NC10 *Rokubacteriales* increased from 1.74% at 1 cm to 3% at 12 cm and decreased from 1.74% at 13 cm to 1.26% at 41 cm. AOV showed a statistical difference in NC10 *Rokubacteriales* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs excluding 3 – 4 (Appendix A.7I). An uncultured *Thermodesulfovibriona* increased gradually with depth from 0.26% at 1 cm to 3.05% at 39 cm. AOV showed a statistical difference in *Thermodesulfovibriona* abundance

across all sample groups and Tukey HSD showed statistical differences between all group pairs except 2 – 3 (Appendix A.7J).

Six ASVs that represented genera and family with specific known catabolic function and that had relatively high abundance as a whole in the core (*i.e.*, < 1%) were targeted. *Desulfobacca* spp. generally increased from 0.1% at 1 cm to 2.54% at 32 cm, and then decreased from 2.07% at 33 cm to 1.55% at 41 cm (Figure 3.17A). AOV showed a statistical difference in *Desulfobacca* spp. abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs excluding 3 – 4 (Appendix A.7K). *Pseudolabrys* spp. 0.38% at 1 cm to 2.07% at 16 cm and had an average abundance of 1.39 ± 0.03% between 17 – 41 cm (Figure 3.17B). AOV showed a statistical difference in *Pseudolabrys* spp. abundance across all groups and Tukey HSD showed statistical differences between group pairs 1 and all other groups and 2 – 4 (Appendix A.7 L). The *Sideroxydans* spp. had an average abundance of 0.16 ± 0.03% between 1 – 22 cm and increased from 0.68% at 23 cm to 5.32% at 38 cm (Figure 3.17C). AOV showed a statistical difference in *Sideroxydans* spp. abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 4 and all other groups (Appendix A.7M). An uncultured *Micropepsaceae* increased from 0.18% at 1 cm to 1.98% at 21 cm, and between 22 – 41 cm had an average abundance of 1.32 ± 0.07% (Figure 3.17D). AOV showed a statistical difference in *Micropepsaceae* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 and all other groups and 2 – 3 (Appendix A.7N). *Nitrospira* spp. increased from 1.37% at 1 cm to 2.57% at 4 cm, decreased to 0.2% at 14 cm, increased to 1.44% at 27 cm, and between 28 – 41 cm had an average abundance of 0.9 ± 0.11% (Figure 3.17E). AOV showed a statistical difference in *Nitrospira* spp. abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 2 – 3 (Appendix A.7O). *Candidatus Nitrosotalea* spp. had an average abundance of 0.03 ± 0.02% between 1 – 9 cm and increased from 0.5% at 10 cm to 3.6% at 18 cm, and then decreased to an average abundance of 1.08 ± 0.07% between 19 – 41 cm (Figure 3.17F). AOV showed a statistical difference in *Candidatus Nitrosotalea* spp. abundance across all sample groups and Tukey HSD showed statistical differences between group pair 1 and all other groups (Appendix A.7P).

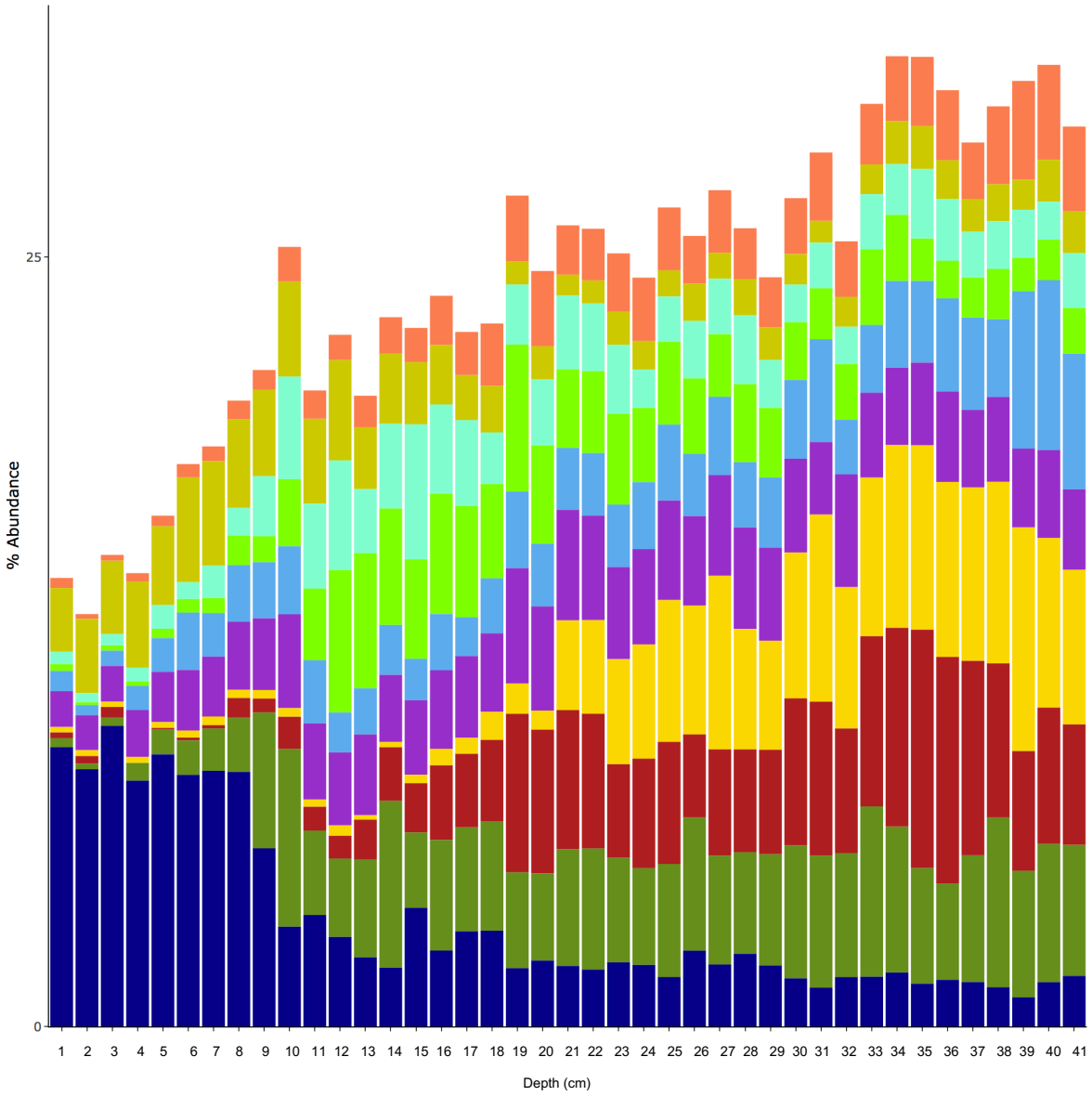


Figure 3.16. Control Core: 100% stacked bar chart of the ten most abundant ASV taxonomies collapsed at genus level. Size of bar denotes percentage abundance of the genera as a proportion of total ASV genera assemblage within each sample. Dark blue = uncultured *Acidobacteria* Subgroup 6; dark green = uncultured methanogenic *Bathyarchaeia*; red = uncultured *Gallionellaceae*; yellow = uncultured α-*proteobacteria*; purple = uncultured *Anaerolineaceae*; blue = uncultured *Nitrospirae* 4-29-1; green = uncultured *Acidobacteriales*; light blue = uncultured δ-*proteobacteria* MBNT15; gold = uncultured *Rokubacteria* NC10; orange = uncultured *Thermodesulfovibriona*.

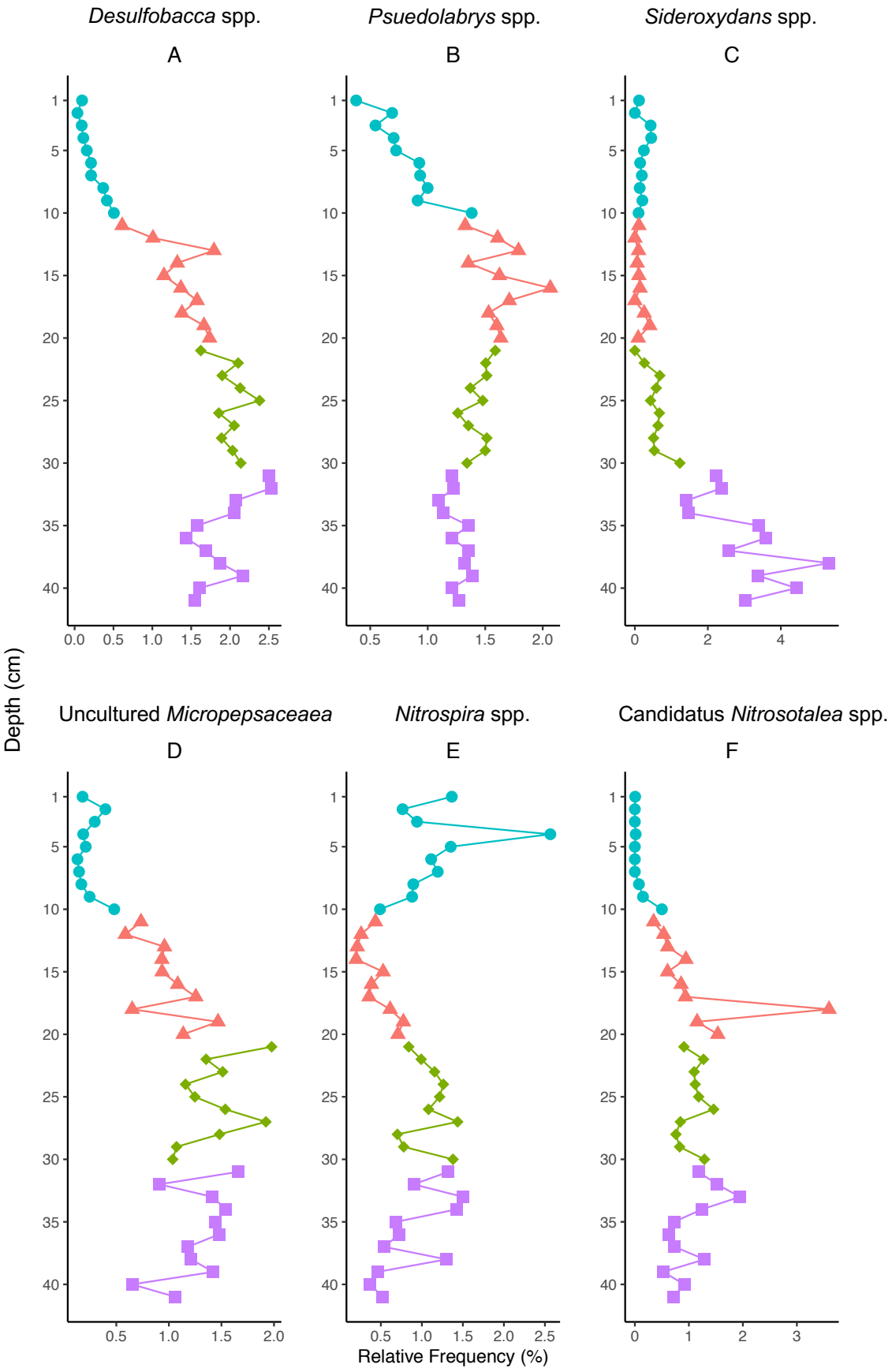


Figure 3.17. Control Core: Depth profiles of six additional key functional microbial taxa. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). *Desulfobacca* spp.; (B). *Pseudolabrys* spp.; (C). *Sideroxydans* spp.; (D). uncultured *Micropepsaceae*; (E). *Nitrospira* spp.; (F). Candidatus *Nitrosotalea* spp.

3.3.5 Tax4Fun Analysis Results

3.3.5.1 *Contaminated Core 'End Point' PCR Target Gene Inferred Abundance Profiles*

The five genes targeted during 'end point' PCR were inferred in the Tax4Fun profiles of the contaminated core. The *amoA* gene (not distinguished between bacterial and archaeal variants in Tax4Fun) had a relatively low abundance compared to the other target genes, and increased from 0.34% at 1 cm to 0.96% at 4 cm, and then between 5 – 23 cm had an average abundance of 0.6 ± 0.02%, and decreased from 0.52% at 24 cm to 0.08% at 34 cm, spiked to 0.63% at 35 cm, and decreased to 0.05% at 37 cm (Figure 3.18A). AOV showed a statistical difference in *amoA* abundance across all sample depth groups and Tukey HSD showed statistical differences between sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.8A). The *nirS* gene increased in abundance moderately from 0.86% at 1 cm to 2.4% at 21 cm, and decreased to 1.03% at 22 cm, and between 23 – 34 cm had an average abundance of 1.24 ± 0.08%, spiked at 35 cm to 4.38%, and between 36 – 41 cm had an average abundance of 0.93 ± 0.09% (Figure 3.18B). AOV and Tukey HSD did not show statistical differences across all or between pairs of sample depth groups for the *nirS* gene (Appendix A.8B). The *nirK* gene persisted with depth, with an abundance of 4.14% at 2 cm to 3.12% at 23 cm, and dropped to 0.54% at 34 cm, spiked to 1.69% at 35 cm, and decreased to 0.43% at 37 cm (Figure 3.18C). AOV showed a statistical difference in *nirK* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.8C). The *dsrA* gene increased from 0.23% at 1 cm to 1.17% at 23 cm and increased from 2.22% at 24 cm to 4.59% at 34 cm, dropped to 2.61% at 35 cm, and between 36 – 41 cm had an average abundance of 4.61 ± 0.17% (Figure 3.18D). AOV showed a statistical difference in *dsrA* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.8D). The *mcrA* gene initially decreased from 1.68% at 2 cm to 0.28% at 8 cm, and increased gradually to 1.96% at 22 cm, and at 25 cm increased to 5.61%, and from 26 – 34 cm had an average abundance of 4.87 ± 0.2%, decreased to 1.07% at 35 cm, increased to 4.81% at 37 cm, and decreased to 2.73% at 41 cm (Figure 3.18E). AOV showed a statistical difference in *mcrA* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.8E).

Inferred gene abundances derived from Tax4Fun and band intensity values of the ‘end point’ PCR target genes were tested to identify any correlation using Spearman correlation. The Tax4Fun *amoA* inferred gene abundances appeared to have a statistical, albeit weak, positive correlation with the archaeal *amoA* PCR band intensities (Figure 3.19A), while there was a stronger positive correlation with the bacterial *amoA* PCR band intensities (Figure 3.19B). The *nirS* gene had a weak positive correlation between methods (Figure 3.19C), and the *nirK* a slightly strong positive correlation (Figure 3.19D). The *dsrA* and *mcrA* genes had the strongest positive correlation between methods (Figures 3.19E and 3.19F).

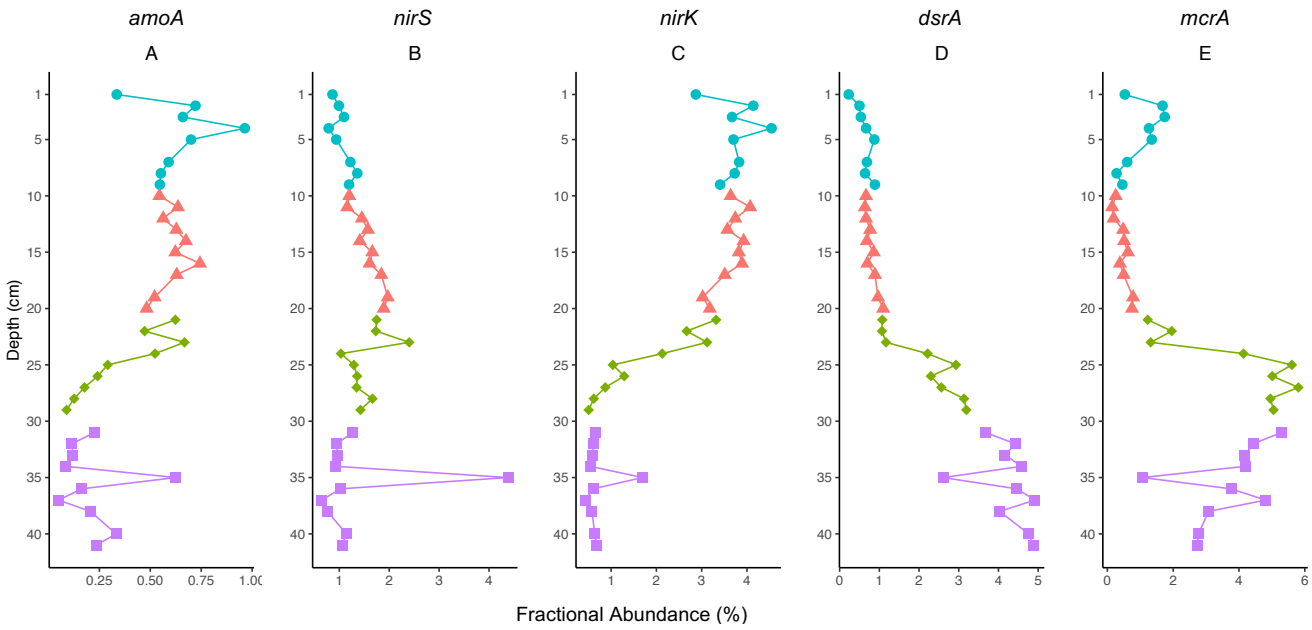


Figure 3.18. Contaminated Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by ‘end point’ PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirS*; (C). *nirK*; (D). *dsrA*; (E). *mcrA*.

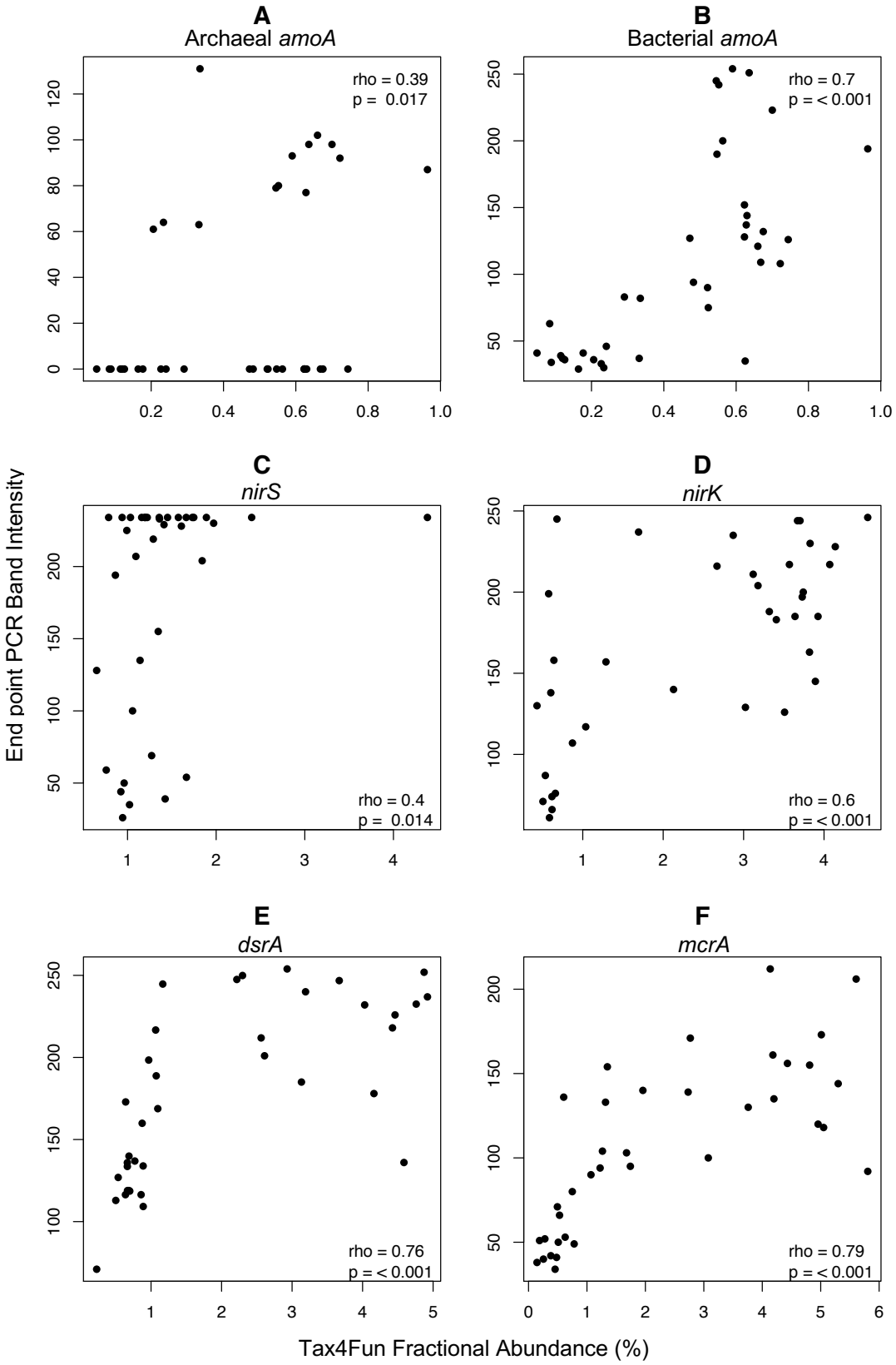


Figure 3.19. Contaminated Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA*.

3.3.5.2 *Contaminated Core Nitrogen Metabolism Inferred Gene Abundance Profiles*

The denitrifying nitrate reductase *narG* persisted with depth, from 7.65% at 1 cm to 6.92% at 23 cm (average abundance of 7.35 ± 0.17%), decreased to 4.13% at 24 cm, and between 25 – 34 cm had an average abundance of 2.48 ± 0.19%), spiked to 11.59% at 35 cm, and decreased to 1.83 ± 0.1% between 36 – 41 cm (Figure 3.20A). AOV showed a statistical difference in *narG* abundance across all sample depth groups and Tukey HSD showed statistical differences between sample depth group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix A.9A). The other denitrifying nitrate reductase *napA* has a very similar distribution pattern to the *nirS* and increased from 0.79% at 1 cm to 1.67% at 23 cm (average abundance of 1.01 ± 0.05%), decreased to 0.86 ± 0.03% between 24 – 34 cm, spiked to 5.7% at 35 cm, and decreased to 0.69 ± 0.05% between 36 – 41 cm (Figure 3.20B). AOV and Tukey did not show statistical differences in *napA* abundances across or between sample group pairs (Appendix A.9B). The nitric oxide reductase *norB* had an identical distribution to the *nirS* and *napA*, albeit with higher values, peaking at 13.71% at 35 cm (Figure 3.20C). AOV and Tukey did not show statistical differences in *norB* abundances across or between sample group pairs (Appendix A.9C). The nitrous-oxide reductase *nosZ* gene was most abundant in the 1 – 23 cm region, with an average abundance of 2.66 ± 0.07%, and decreased to 1.07 ± 0.08% between 24 – 41 cm (Figure 3.20D). The nitrite reductase *nrfA* from dissimilatory nitrate reduction increased in abundance with depth, with an average abundance of 1.26 ± 0.04% between 1 – 23 cm and an average of 3.35 ± 0.22% between 24 – 41 cm, including a decrease to 2.05% at 35 cm (Figure 3.20E). AOV showed a statistical difference in *nrfA* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.9E). From the nitrogen fixation pathway, the nitrogenase *nifD* gene, had an identical pattern of distribution to the *nrfA* albeit with higher abundances (Figure 3.20F), and AOV showed a statistical difference in *nifD* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.9F).

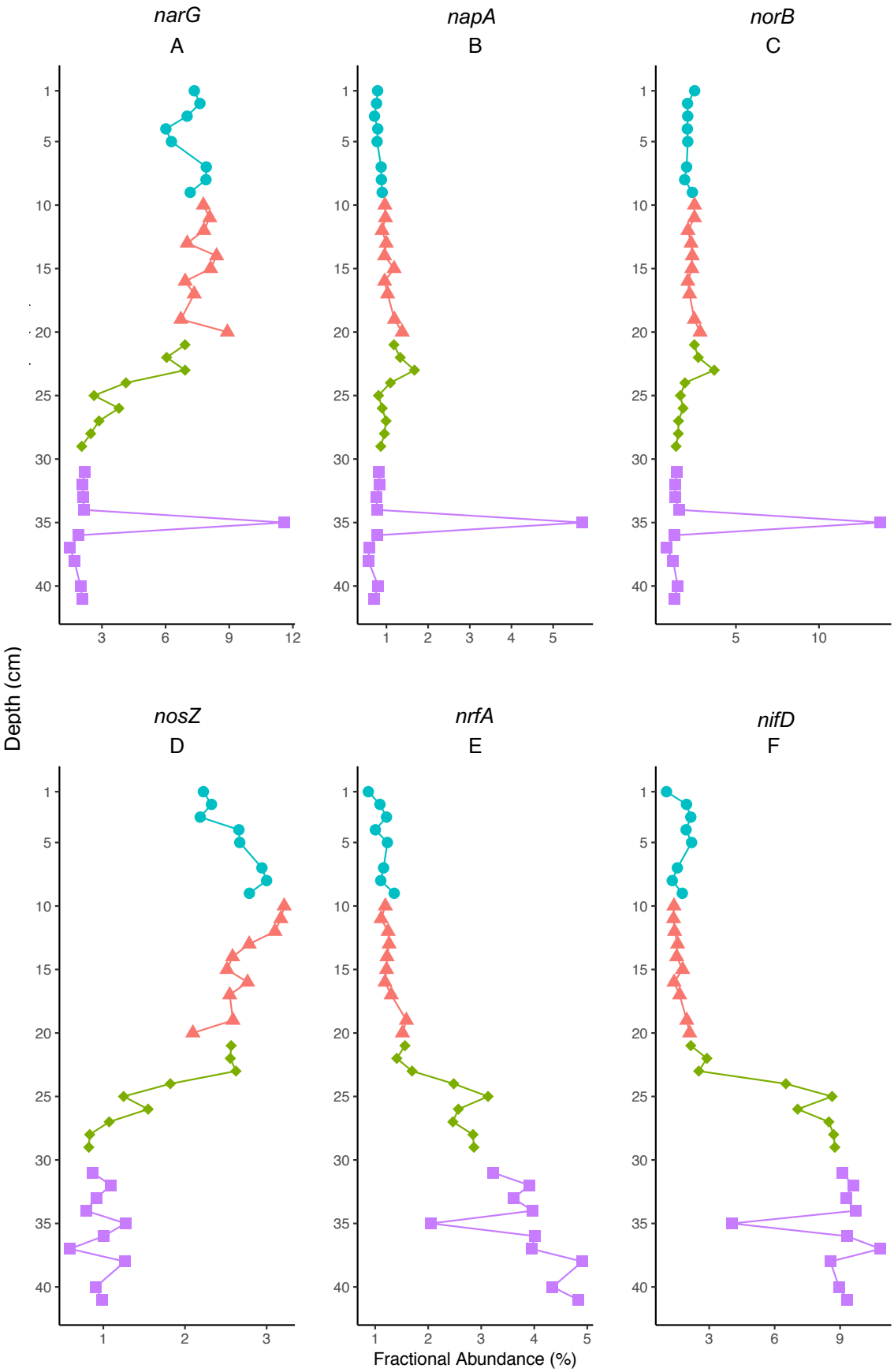


Figure 3.20. Contaminated Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD*.

3.3.5.3 *Contaminated Core Profiles of Inferred Abundances of Additional Target Genes*

A gene from the anaerobic carbon fixating Wood-Ljungdahl pathway, the acetyl-CoA synthase *acsB* gene, followed a similar pattern of distribution to the *dsrA* gene, having a steadily low abundance between 1 – 23 cm of 0.91 ± 0.06%, which increased to 5.45 ± 0.36% between 24 – 34 cm, decreased to 3.02% at 35 cm, and increased to 6.72 ± 0.31% between 36 – 41 cm (Figure 3.21A). AOV showed a statistical difference in *acsB* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.10A). The adenylylsulfate reductase *aprA* gene from the dissimilatory sulfate reductase pathway followed a similar distribution to the *dsrA* and *acsB*: 1.39 ± 0.1% between 1 – 23 cm; 6.36 ± 0.52% between 24 – 34 cm; 4.8% at 35 cm; 8.74 ± 0.3% between 36 – 41 cm (Figure 3.21 B). AOV showed a statistical difference in *aprA* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.10B). The methane monooxygenase *mmoX* gene had an abundance of 0.24% at 1 cm and 0.15 ± 0.003% between 2 – 34 cm, 0.22% at 35 cm, and 0.11 ± 0.006% between 36 – 41 cm (Figure 3.21C). AOV showed a statistical difference in *mmoX* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 4 and 2 – 4 (Appendix A.10C). The methanol dehydrogenase *mxaF* gene had an average abundance of 0.51 ± 0.03% between 1 – 34 cm, increased to 4.84% at 35 cm, and decreased to 0.35 ± 0.03% between 36 – 41 cm (Figure 3.21D). AOV and Tukey did not show statistical differences in *mxaF* abundances across or between sample group (Appendix A.10D). The NiFe hydrogenase *hydB* gene generally had a low relative abundance, not exceeding 0.07%, and had an abundance of 0.033 ± 0.001% between 1 – 23 cm, and 0.058 ± 0.002% between 24 – 41 cm (Figure 3.21E). AOV showed a statistical difference in *hydB* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.10E). The methanogenic Wolfe cycle methylene-tetrahydromethanopterin dehydrogenase *mtdB* gene had a similar distribution to the *napA*, *nirS*, *norB*, and *mxaF* in that it had relatively low abundance throughout except for a peak at 35 cm (Figure 3.21F). AOV and Tukey did not show statistical differences in *mtdB* abundances across or between sample depth groups (Appendix A.10F). The methanogenic heterodisulfide reductase *hdrA* gene had a similar pattern

of distribution as the *acsB*, *aprA*, *dsrA*, *mcrA*, with an average abundance of 6.76 ± 0.46% between 1 – 23 cm, 34.51 ± 1.94% between 24 – 34 cm, 16.93% at 35 cm, and 38.34 ± 2% between 36 – 41 cm (Figure 3.21G). AOV showed a statistical difference in *hdrA* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.10G). The anaerobic carbon fixation Wood-Ljungdahl pathway, anaerobic carbon-monoxide dehydrogenase *cooS* gene followed the same pattern as *acsB*, *dsrA*, *mcrA* and similar genes, with average abundances of 1.21 ± 0.12% between 1 – 23 cm, 7.45 ± 0.47% between 24 – 34 cm, 2.6% at 35 cm, and 7.42 ± 0.75% between 36 – 41 cm (Figure 3.21H). AOV showed a statistical difference in *cooS* abundance across all sample groups and Tukey HSD showed statistical differences between all group pairs except 1 – 2 (Appendix A.10H).

3.3.5.4 *Control Core 'End Point' PCR Target Gene Inferred Abundance Profiles*

The five genes targeted during 'end point' PCR were inferred in the Tax4Fun profiles of the control core. The *amoA* gene initially decreased from 0.47% at 1 cm to 0.11% at 13 cm, increased gradually to a peak of 0.7% at 35 cm, and decreased to 0.43% at 41 cm (Figure 3.22A). AOV showed a statistical difference in *amoA* abundance across all sample depth groups and Tukey HSD showed statistical differences between all pairs of sample depth groups, excluding pair 1 – 3 (Appendix A.11A). The *nirS* had a similar distribution to the *amoA*, decreasing from 1.32% at 3 cm to 0.53% at 22 cm, increased again to 1.16% at 36 cm, and decreased to 0.92% at 41 cm (Figure 3.22B). AOV showed a statistical difference in *nirS* abundance across sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2, 1 – 3, 2 – 4, and 3 – 4 (Appendix A.11B). The *nirK* gene was most abundant in the upper 10 cm region, decreasing from 4.75% at 4 cm to 1.79% at 11 cm, and between 12 – 41 cm, had an average abundance of 1.2 ± 0.04% (Figure 3.22C). AOV showed a statistical difference in *nirK* abundance across sample groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix A.11C). The *dsrA* gene increased gradually with depth, from 0.27% at 2 cm to 4.42% at 39 cm (Figure 3.22D). AOV showed a statistical difference in *dsrA* abundance across all sample groups and Tukey HSD showed statistical differences between all pairs of sample groups, except pair 2 – 3 (Appendix A.11D). The *mcrA*

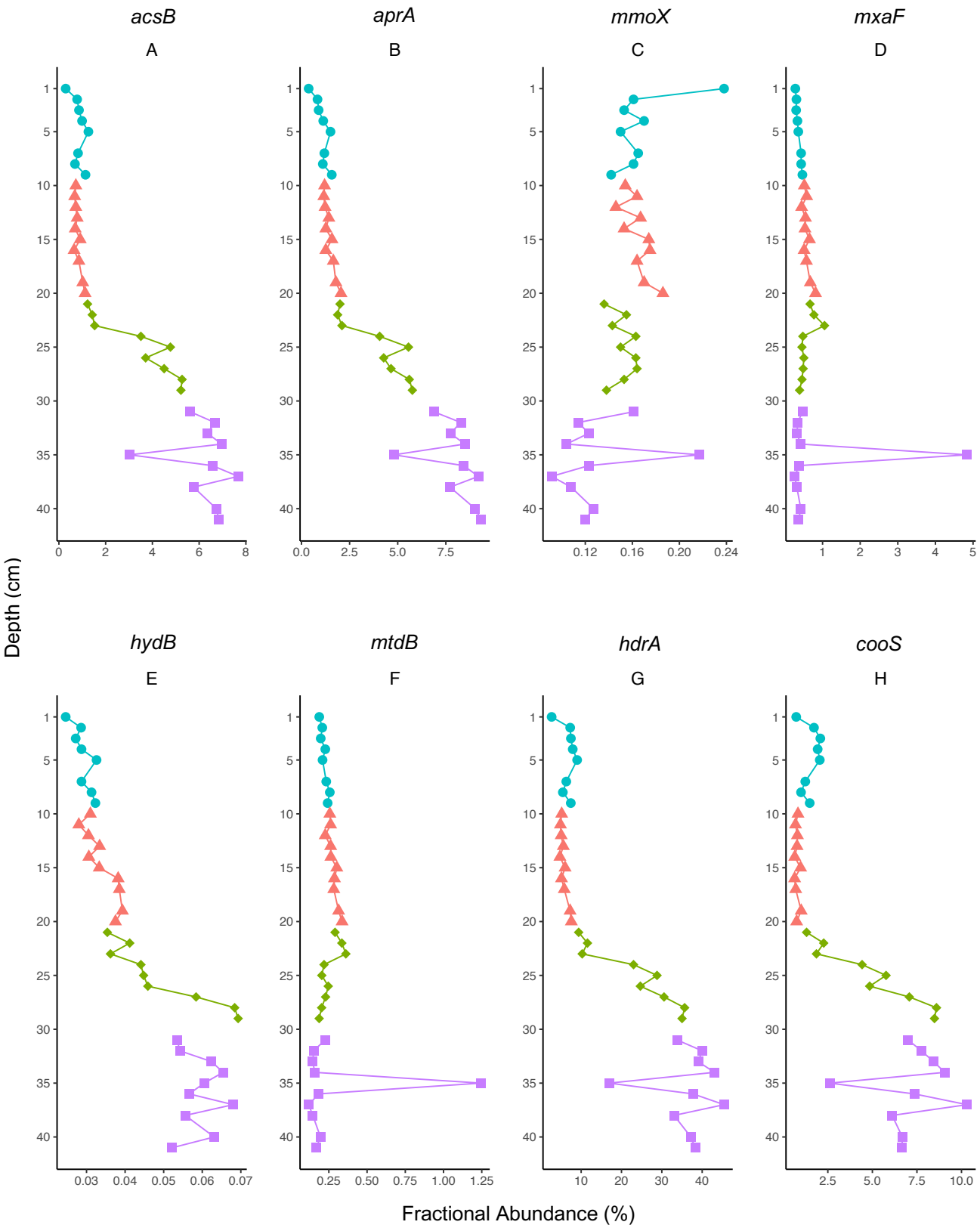


Figure 3.21. Contaminated Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS*.

gene had very sporadic inferred distribution, including many zero counts throughout the core. There was no pattern with depth (Figure 3.22E). AOV showed a statistical

difference in *mcrA* abundance and Tukey HSD showed a statistical difference between group pair 1 – 2 (Appendix A.11E).

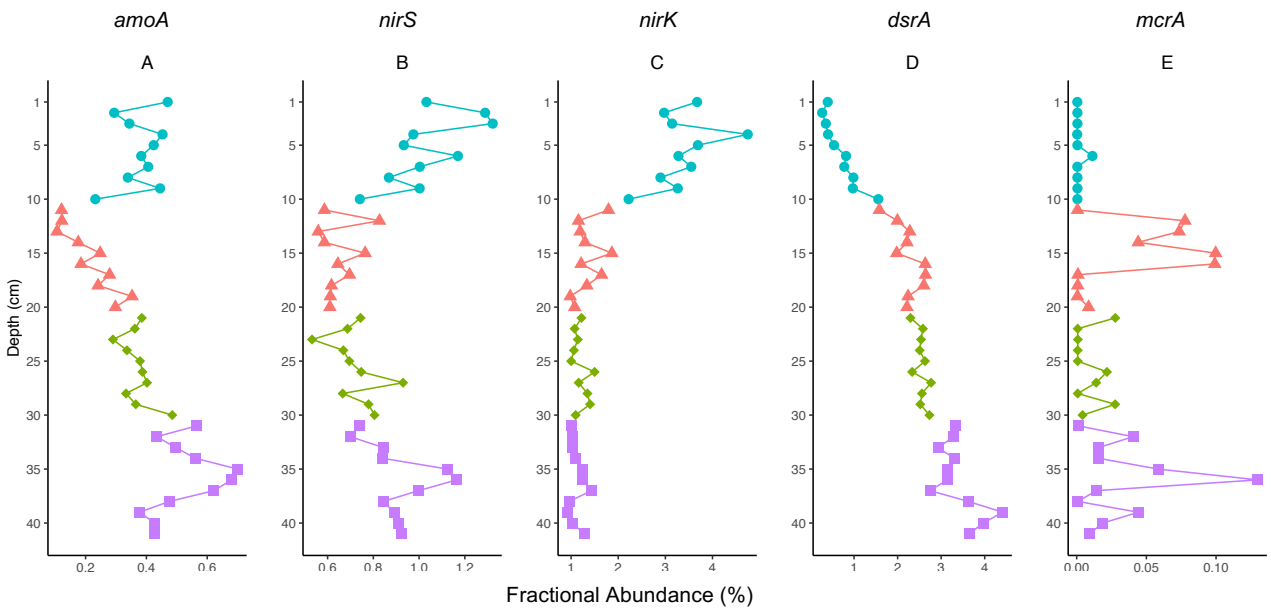


Figure 3.22. Control Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirS*; (C). *nirK*; (D). *dsrA*; (E). *mcrA*.

Inferred abundances of the 'end point' PCR target genes from Tax4Fun were correlated with the band intensities of the same genes from 'end point' PCR. The Tax4Fun *amoA* profile did not correlate with either bacterial or archaeal 'end point' PCR band intensities (Figures 3.23A and 3.23B). The *nirS* gene did not correlate (Figure 3.23C), but the *nirK* appeared to show a weak, albeit statistically positive correlation between abundance profiles (Figure 3.23D). The *dsrA* visually appeared to have a positive correlation between abundance profiles, but this was not reflected in the Spearman correlation values (Figure 3.23E). Peculiarly, the *mcrA* gene did appear to have a positive correlation in abundances from the two profiles (Figure 3.23F).

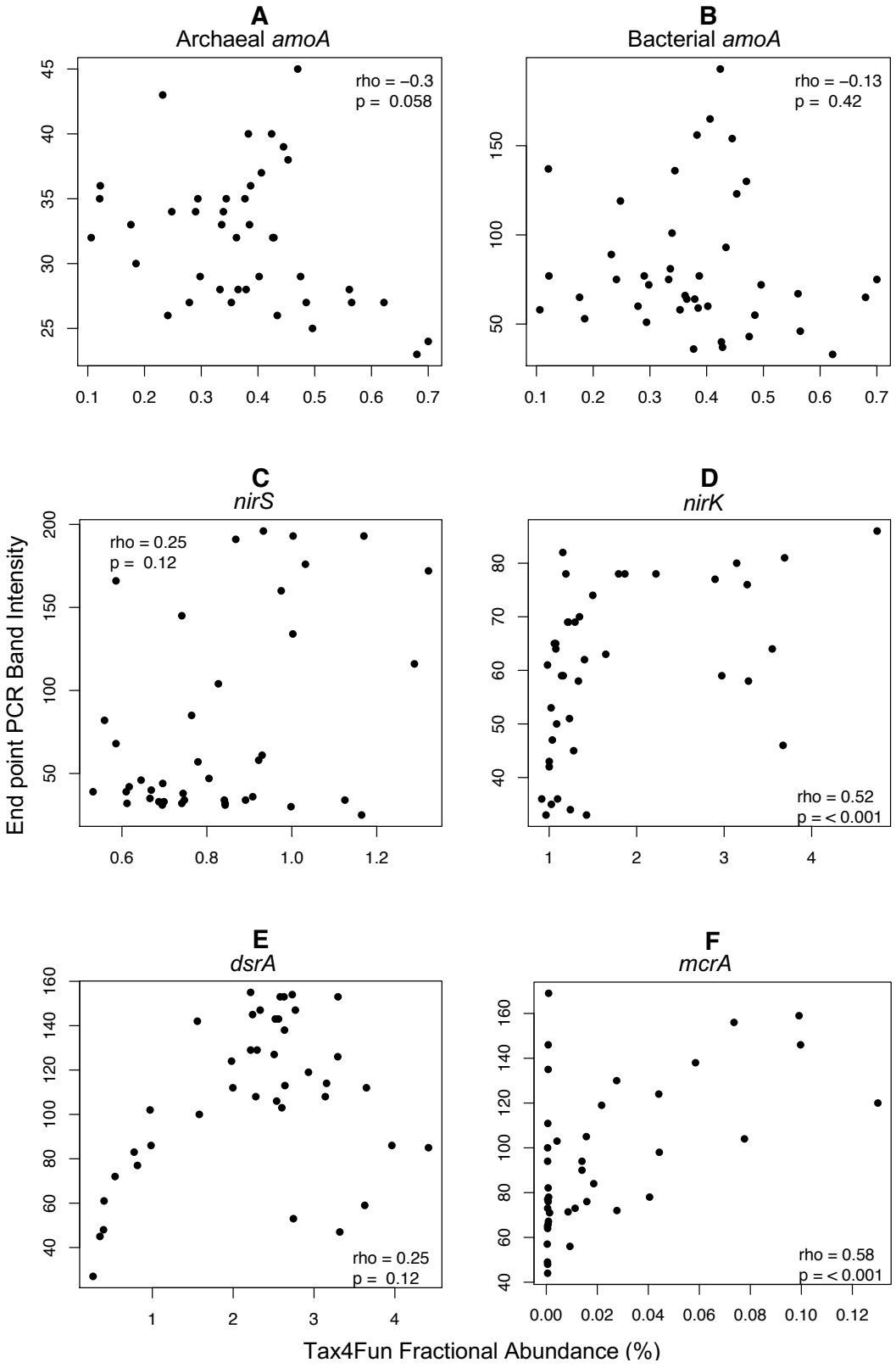


Figure 3.23. Control Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirS*; (D). *nirK*; (E). *dsrA*; (F). *mcrA*.

3.3.5.5 *Control Core Nitrogen Metabolism Inferred Gene Abundance Profiles*

The *narG* was most abundant in the upper 10 cm of the core, decreasing from 14.09% at 3 cm to 2.24% at 19 cm, and between 20 – 41 cm, had an average abundance of 2.82 ± 0.08% (Figure 3.24A). AOV showed a statistical difference in *narG* abundance across all sample depth groups, and Tukey HSD showed statistical differences between sample group 1 and all other groups (Appendix A.12A). The *napA* appeared to have little depth-related distribution of abundance, and between 1 – 38 cm, had an average abundance of 0.88 ± 0.01%, dropping to 0.69% at 41 cm (Figure 3.24B). AOV and Tukey did not show statistical differences across or between sample depth groups (Appendix A.12B). The *norB* gene had a similar pattern to the *napA* (Figure 3.24C), although AOV showed a statistical difference in *norB* abundance across all sample groups and Tukey HSD showed a statistical difference between group pair 1 – 2 (Appendix A.12C). The *nosZ* was most abundant from 1 – 11 cm, with an average abundance of 2.59 ± 0.08%, which decreased to 1.5 ± 0.03% between 12 – 41 cm (Figure 3.24D). AOV showed a statistical difference in *nosZ* abundance across all sample depth groups, and Tukey HSD showed statistical differences between sample group 1 and all other groups (Appendix A.12D). The *nrfA* gene increased with depth from 1.07% at 2 cm to 2.65% at 18 cm, decreased to 1.88% at 19 cm, and increased to 3.83% at 39 cm (Figure 3.24E). AOV showed a statistical difference in *nrfA* abundance across all sample depth groups, and Tukey HSD showed statistical differences between all pairs of groups except 2 – 3 (Appendix A.12E). The *nifD* increased in a similar pattern to the *nrfA*, from 1.08% at 2 cm to 4.2% at 18 cm, decreased to 3.53% at 19 cm, and increased to 6.71% at 39 cm (Figure 3.24F). AOV showed a statistical difference in *nifD* abundance across all sample depth groups, and Tukey HSD showed statistical differences between all pairs of groups except 2 – 3 (Appendix A.12F).

3.3.5.6 *Control Core Profiles of Inferred Abundances of Additional Target Genes*

The *acsB* gene increased gradually with depth from 0.28% at 2 cm to 6.35% at 39 cm (Figure 3.25A). AOV showed a statistical difference in *acsB* abundance across all sample depth groups, and Tukey HSD showed statistical differences between all pairs of groups except 2 – 3 (Appendix A.13A). The *aprA* gene also

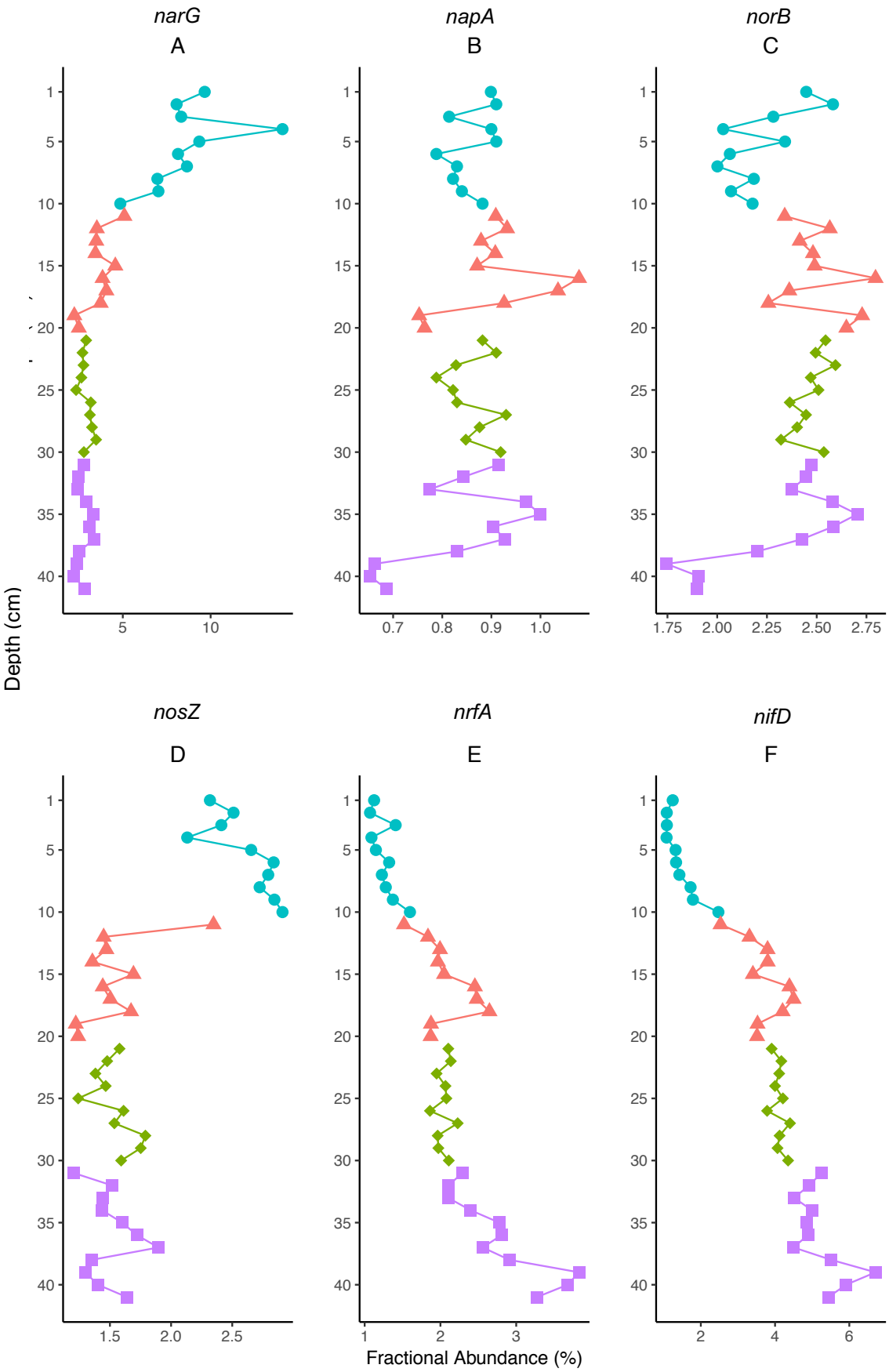


Figure 3.24. Control Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD*.

increased gradually with depth, from 0.45% at 2 cm to 7.97% at 39 cm (Figure 3.25B). AOV showed a statistical difference in *aprA* abundance across all sample depth groups, and Tukey HSD showed statistical differences between all pairs of groups except 2 – 3 (Appendix A.13B). The *mmoX* gene decreased with depth from 0.26% at 2 cm to 0.09% at 19 cm, and between 20 – 41 cm had an average abundance of 0.11 ± 0.003% (Figure 3.25C). AOV showed a statistical difference in *mmoX* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix A.13C). The *mxaF* gene generally increased with depth, from 0.25% at 4 cm to 0.55% at 35 cm and decreased to 0.31% at 41 cm (Figure 3.25D). AOV showed a statistical difference in *mxaF* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group 4 and all other groups (Appendix A.13D). The *hydB* gene increased gradually with depth, from 0.019% at 3 cm to 0.098% at 31 and decreased to 0.074% at 41 cm (Figure 3.25E). AOV showed a statistical difference in *hydB* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group all group pairs (Appendix A.13E). The *mtdB* gene initially decreased from 0.19 ± 0.005% between 1 – 12 cm, increased from 0.14% at 13 cm to 0.25% at 36 cm, and decreased to 0.16% at 41 cm (Figure 3.25F). AOV showed a statistical difference in *mtdB* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 2 and 2 – 4 (Appendix A.13F). The *hdrA* gene increased gradually with depth from 2.59% at 2 cm to 42.62% at 39 cm (Figure 3.25G). AOV showed a statistical difference in *hdrA* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group all group pairs except 2 - 3 (Appendix A.13G). The *cooS* gene also increased with depth, from 0.41% at 2 cm to 8.61% at 32 cm and decreased to 6.6% at 41 cm (Figure 3.25H). AOV showed a statistical difference in *cooS* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group all group pairs (Appendix A.13H).

3.3.6 Geochemical Analysis Results

3.3.6.1 *XRF Bulk Biogeochemistry of the contaminated core*

When dried, the soil was up to 85 % organic matter at the surface and this decreased steadily with depth to a minimum of ~ 25 wt % at 30 cm (Figure 3.26A; red). In contrast the soil Si content increased from 1 wt. % at the surface to 22 wt %

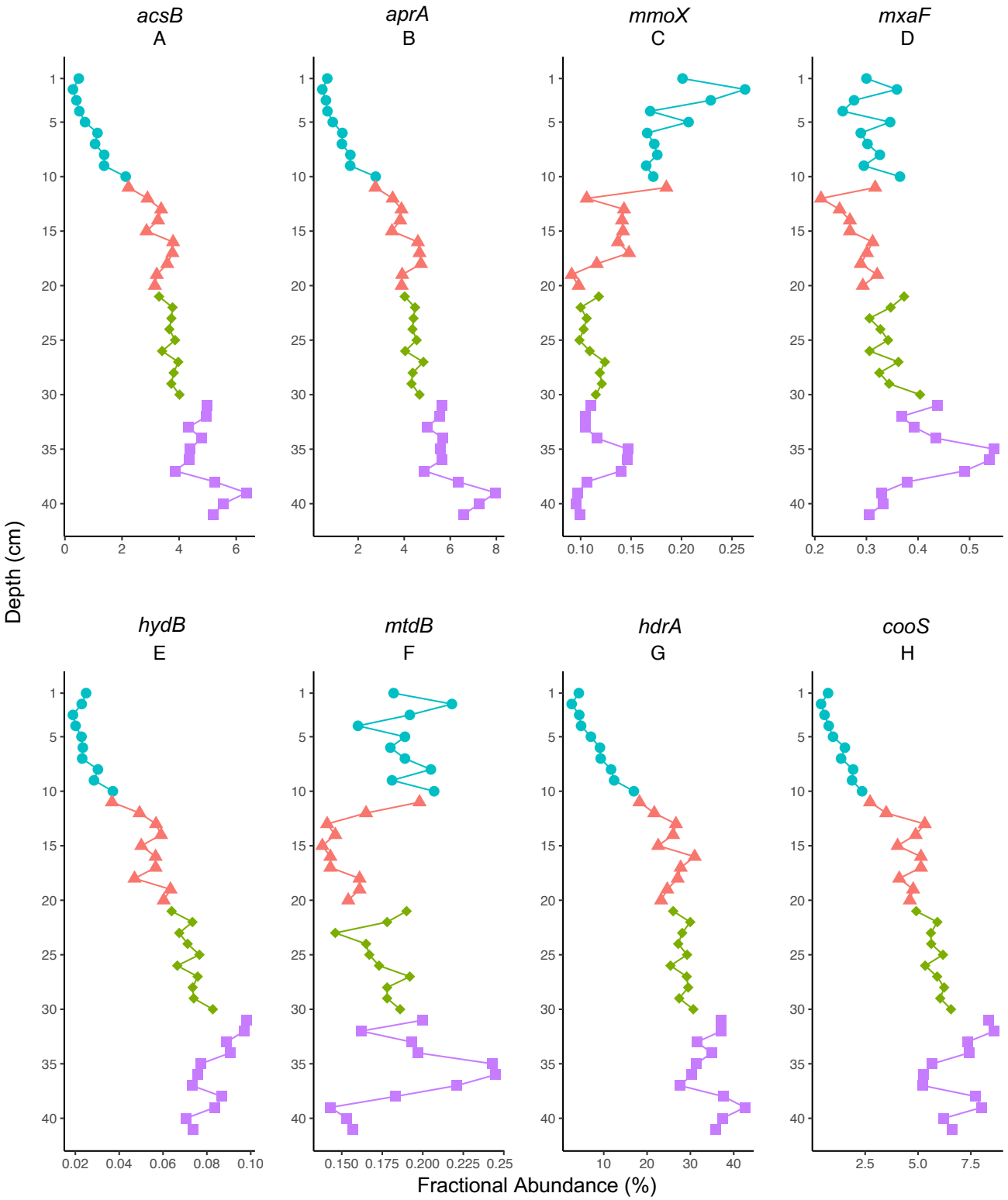


Figure 3.25. Control Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). NiFe hydrogenase; (F). *mtdB*; (G). *hdrA*; (H). *cooS*.

at 30 cm, reflecting an increasing mineral fraction with depth and suggesting that the decrease in C largely reflects dilution (Figure 3.26A; blue). Manganese and Fe were normalised to Al to resolve any redox-driven variability with depth (Morford and

Emerson, 1999). There was a decrease in Mn/Al between 0 and 12 cm (Figure 3.26B). Solid phase Fe (Figure 3.26C) showed a similar trend although the Fe/Al ratio was stable down to 10 cm. Around 15 cm there was a roughly 2/3rd decrease in the Fe/Al ratio, indicative of Fe^{3+} reduction (Tribovillard *et al.*, 2006). The Fe/Al ratio then remained stable until around 23 cm, when it then fell to a minimum at 30 cm. This trend was mirrored by a concurrent increase in porewater Fe (Figure 3.26F) between 15 and 25 cm. Ferrozine analysis of the 0.5 N HCl extracted solid phase Fe (Figure 3.26E) also confirmed that extractable Fe was present as ~ 100 % Fe^{2+} throughout the sediment except for a region between ~ 5 – 15 cm where up to 40 % was present as Fe^{3+}. The increase in 0.5 N HCl extractable Fe^{2+} below 15 cm is consistent with the decrease in solid phase Fe/Al (Figure 3.26C).

Uranium concentrations throughout the core were enriched above natural background concentrations (2.76 mg kg^{-1} (Herring, 2013)) with a maximum of ~1600 mg kg^{-1} at 8 – 10 cm (Figure 3.26D). Below these the solid-phase U concentrations gradually decreased to a minimum of 47 mg kg^{-1} at 28 cm. Analysis of groundwater showed the highest aqueous U concentration around 20 cm (0.19 mg L^{-1}, Figure 3.26G).

3.3.6.2 *Porewater Metal Analysis*

AOV showed a statistical difference U porewater concentrations across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 - 4 (Appendix A.14A). Manganese (Mn) porewater concentrations were highest at 1 cm, at 295.4 ppm, and decreased to 127.2 ppm at 12 cm, and to 75.8 ppm at 25 cm. Between 26 – 36 cm, the average concentration was 81.25 ± 1 ppm, and increased to 94.08 ± 2.78 ppm between 37 – 41 cm (Appendix B.1A). AOV showed a statistical difference Mn porewater concentrations across all sample depth groups, and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix A.14A). Porewater zinc (Zn) concentrations increased from 790.6 ppm at 1 cm to a peak of 2043.2 ppm at 8 cm, decreased gradually to 1072.2 ppm at 24 cm, and decreased further to 55.8 ppm at 33 cm. There was an increase to 323.3 ppm at 35 cm, which decreased to 83 ppm at 41 cm (Appendix B.1B). AOV showed a statistical difference Zn porewater concentrations across all sample depth groups, and Tukey HSD showed statistical differences between all pairs of groups except 1 – 2 (Appendix A.14C).

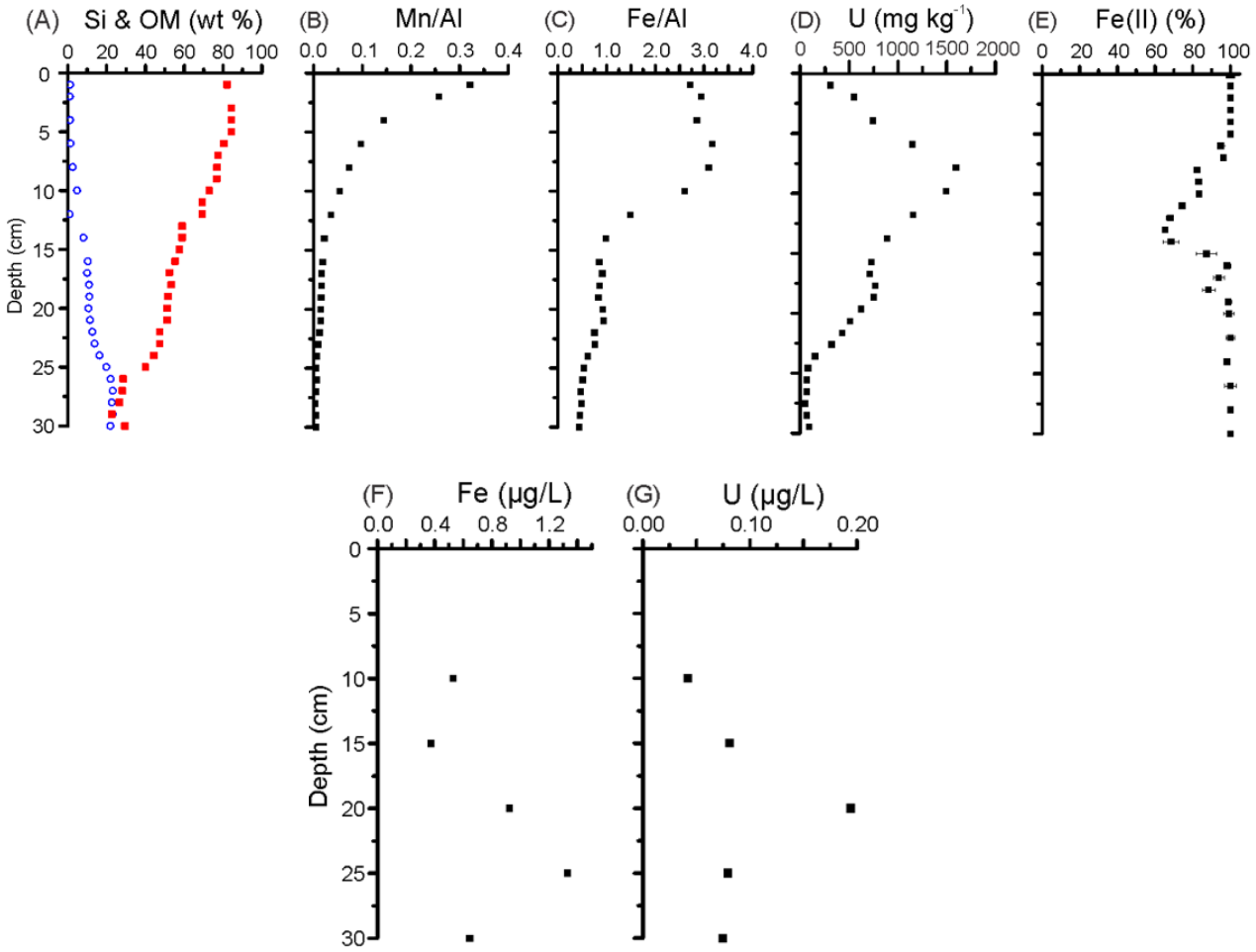


Figure 3.26. Contaminated Core: From Fuller *et al.* (*in prep.*). Depth profiles of: (A) Si = ○, organic matter (OM) = ■; (B) Mn/Al ratio; (C) Fe/Al ratio; (D) U concentration; and (E) HCl extractable solid phase Fe^{2+} as a % of total extractable Fe. Porewater: (F) Fe, and (G) U concentrations. From Fuller *et al*. (in prep.).

3.3.6.3 *Uranium L3-edge XANES and M4-edge HERFD-XANES*

The solubility, and hence mobility of U is largely controlled by its oxidation state. At Needle's Eye, U concentrations were highest at 8 – 10 cm depth (Figure 3.26D). Overall, the XANES shows that the accumulation of U in the upper section of the core (above 23 cm; principally between 10 and 15 cm) is not occurring *via* reductive precipitation. U^{6+} persists as the main oxidation state as deep as 23 cm (Appendix B.2A; Table 3.4), despite the evidence of Fe^{3+} reduction. Only at 23 cm does there appear to be a reduction of ~ 20 % of the U^{6+} to U^{4+} (Table 3.3) (Figure 3.27). Furthermore, even under these anoxic conditions, around a third of the U remains as U(VI) (see Appendix B.2B where U^{6+} peaks are still evident at 30 cm).

Table 3.3. Contaminated Core: Uranium oxidation state variation with sample depth as determined by LCF fitting of the U L_3-edge XANES and ITFA analysis of the U M_4-edge HERFD-XANES. All percentages are normalised to 100.

Depth (cm)	% U(IV)	% U(VI)
	L_3	
1	22	78
5	37	63
10	8	92
15	15	85
20	10	90
23	20	80
25	42	58
27	57	43
30	90	10
35	73	27
	M_4	
10	25	75
20	20	80
25	40	60
30	70	30
35	25	75

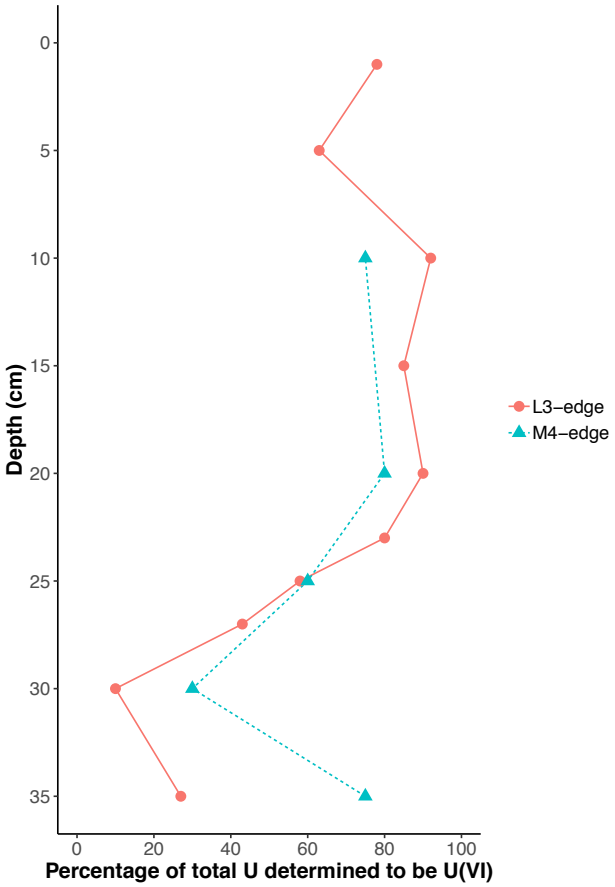


Figure 3.27. Contaminated Core: Depth profile of percentage of total U determined to be U^{6+} versus U^{4+} by LCF fitting of the U L_3-edge XANES (red line, circle points) and ITFA analysis of the U M_4-edge HERFD-XANES (blue dotted line, triangles).

3.3.7 Environmental Fitting

Non-metric multidimensional scaling (nMDS) based on Bray Curtis dissimilarity and detrended correspondence analysis (DCA) plots were produced as above with the addition of environmental variable fitting using the porewater metal concentrations and TOC values for the contaminated core, to identify any potential correlations between the geochemical variables and the nMDS and DCA axes, and how this corresponds to depth sample distribution on the same plots. The nMDS with environmental fitting appeared to show TOC values corresponding with the upper surface samples (2, 3 cm), the U and the Zn corresponding with the depth samples of 15 – 17 cm, and the Mn correlating with nMDS axes in such a fashion that did not correspond with distribution of the depth samples (Figure 3.28). The DCA with the environmental fitting appeared to show TOC, Mn, and Zn all corresponding with the surface samples, and U corresponding with the samples on the DCA1 axis, but less so on the DCA2 axis (Figure 3.29).

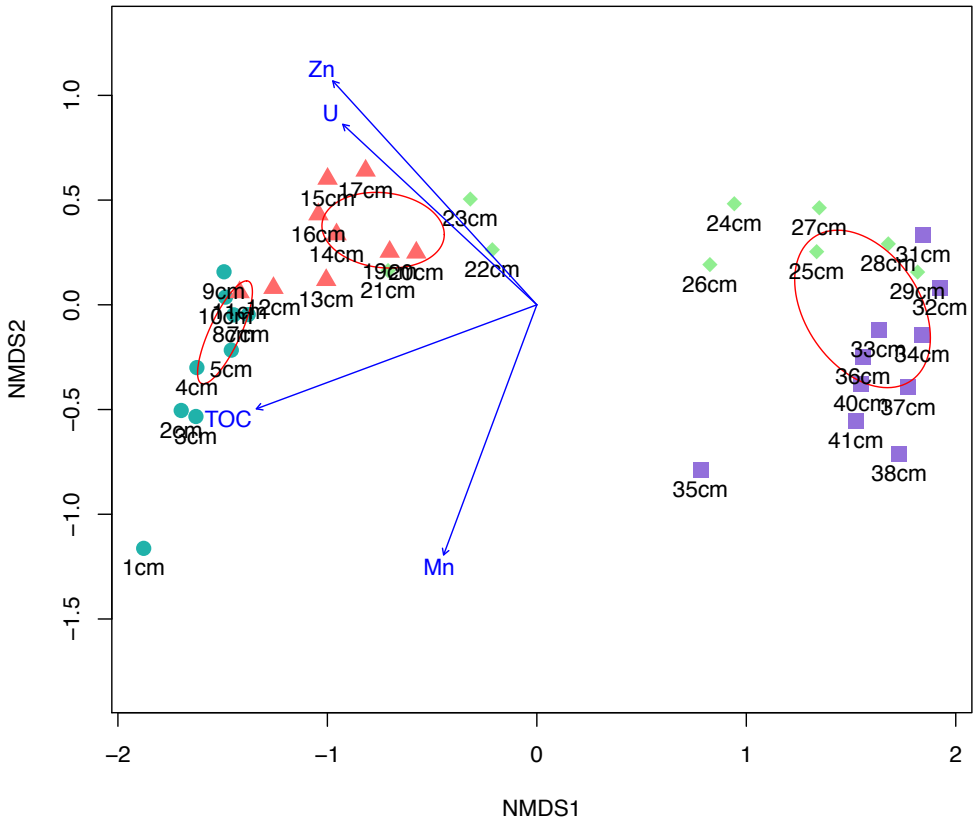


Figure 3.28. Contaminated Core: nMDS based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Environmental Fitting correlation is represented by the blue arrows. Colours and shapes denote sample depth group.

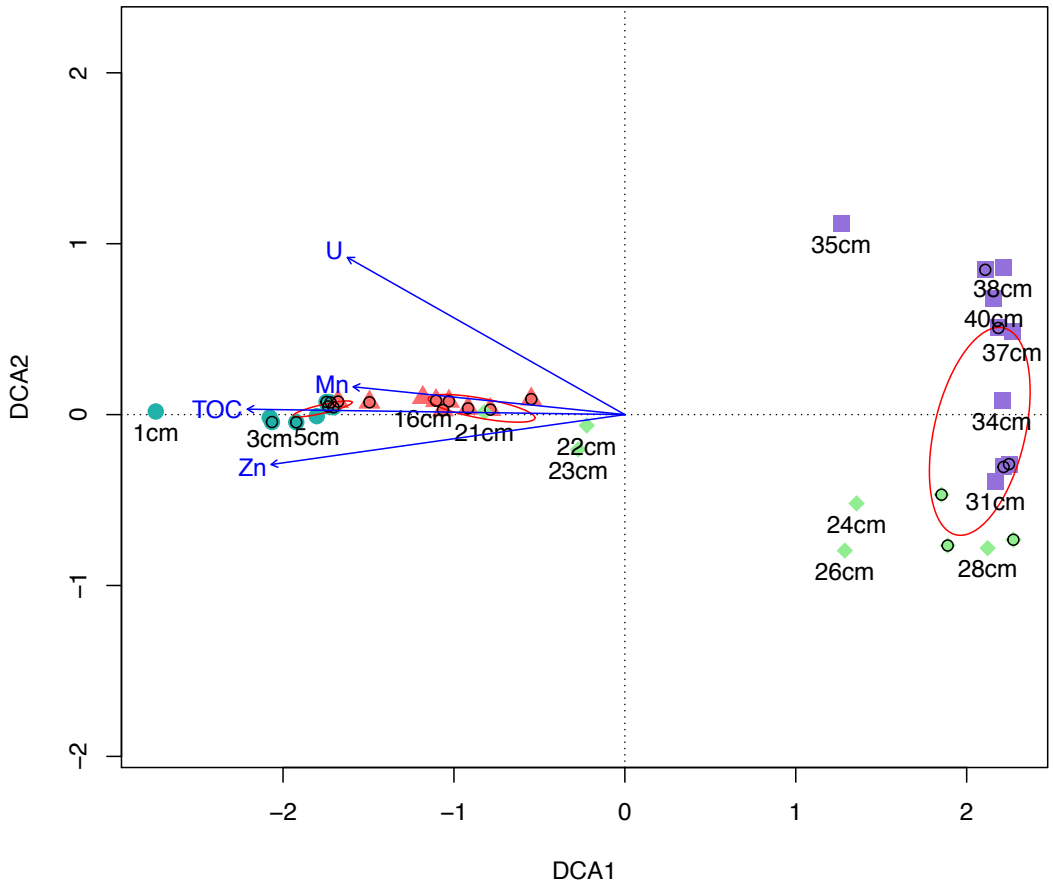


Figure 3.29. Contaminated Core: DCA of the microbial community assemblages between samples. Environmental fitting correlation is represented by the blue arrows.

3.3.8 GraftM Analysis Results

Raw sequences from shotgun metagenome sequencing were input and searched against a range of GraftM functional gene reference packages. Sequences were matched against the reference package using Hidden Markov Models, classified microbial taxonomy or a gene group classification. The results from the GraftM analysis using a range of functional gene reference packages are presented below. Abundances of classified gene taxonomies or groups represent the normalised percentage abundance of that particular group as a function of all the sequences identified for that particular gene target as a function of that particular depth sample.

3.3.8.1 *16S rRNA Gene Taxonomy GraftM Results*

The relative abundance of 16S rRNA gene sequence copies identified by GraftM as a proportion of all the sequences in the six samples submitted for shotgun metagenome sequencing were: 0.038% of all reads at 1 cm; 0.043% for 3 cm; 0.043% for 10 cm; 0.047% for 17 cm; 0.054% for 23 cm, and; 0.048% for 31 cm. Specifically, archaeal sequences increased from 5% of all 16S rRNA gene reads identified at 1 cm to 12% at 31 cm. The ten most abundant sequences clustered at genus level were identified. An uncultured β-*proteobacteria* group was the most abundant single classification throughout the core and peaked in abundance at 9 cm (5.56%). An *Acidobacteria* 6 group was the second most abundant and had the highest abundance also at 9 cm (5.2%) and by 31 cm was very low in abundance (0.08%). An uncultured *Proteobacteria* appeared to increase from 1 (0.28%) to 23 cm (4.56%) and decrease at 31 cm (1.85%). The *Gallionellaceae* appeared to decrease with depth consistently, from 3.19% at 1 cm to 0.08% at 31 cm. The *Chloroflexi* Ellin6529 peaked in abundance at 17 cm (2.63%). The *Nitrospirales* increased from 1.56% at 1 cm to 3.4% at 17 cm and then decreased to 0% at 31 cm. The δ-*proteobacteria* family *Syntrophobacteraceae* also increased from 1 cm (0.99%) to 17 cm (3.29%) and decreased by 31 cm (1.76%). The δ-*proteobacteria* MBNT15 increased from 1 cm (0%) to 17 cm (3.62%) and decreased slightly to 2.89% at 31 cm. The δ-*proteobacteria* BPC076 increased from 0% at 1 cm to 8.09% at 31 cm. The candidate phylum *Parcubacteria* (Nelson and Stegen, 2015) increased from 1 cm (0.64%) to 17 cm (2.85%) and decreased to 31 cm (0.88%) (Appendix C.1).

Although not in the profile of the ten most abundant taxa, archaeal classification of 16S rRNA gene identified sequences increased from 2% of identified sequences at 1 cm to 13% at 31 cm. The *Euryarchaetoa*, including the *Methanosarcina* and *Methanosaeta*, increased from 18% of archaea at 1 cm to 38% at 31 cm. The *Thaumarchaeota* decreased from 3% of archaea at 1 cm to 0% at 31 cm. Additionally, the δ-*proteobacteria* increased from 7% of identified reads at 1 cm to 18% at 31 cm. In addition to the *Syntrophobacteraceae*, BPC076, and MBNT15 groups in Appendix C.1, the *Desulfobacca* also represented 1% of identified sequences at 31 cm.

3.3.8.2 *Nitrate reductase narG GraftM Results*

The relative abundance of *narG* gene sequence copies identified by GraftM increased as a proportion of the total sequences in each sample with depth from 0.11% of sequences at 1 cm to 0.24% at 31 cm. Taxonomic assignments of the *narG* fragments also changed with depth. For instance, *narG* assigned to the family *Desulfobacteraceae* (was the most abundant taxonomic group), and increased from 1.82% at 1 cm to 13.77% at 31 cm. The *Actinobacteria* genus *Pseudonocardia* decreased from 6.04% at 1 cm to 1.77% at 31 cm. An *Actinobacteria* assigned *narG* gene decreased from 6.62% at 1 cm 2.18% at 31 cm. A *Proteobacteria* decreased from 5.06% at 1 cm to 2.44% at 31 cm. The *Actinobacteria* genus *Saccharopolyspora* decreased from 5.15% at 1 cm to 0.22% at 31 cm. The δ-*proteobacteria* family *Syntrophorhabdaceae* increased from 0.18% at 1 cm to 9.85% at 31 cm. An α-*proteobacteria* increased from 2.75% at 1 cm to 3.27% at 17 cm and decreased to 1.71% at 31 cm. The *Planctomycetes* genus *Kuenia* peaked at 2.92% at 3 cm, and had a minimum of 2.08% at 17 cm. The *Anaerolineae* increased from 0.89% at 1 cm to 2.83% at 17 cm and decreased to 2.22% at 31 cm (Appendix C.2).

Although archaeal taxonomic assigned *narG* sequence copies never exceeded 5% of the total *narG* fragment community (and thus did not feature in the profile of the ten most abundant taxonomic classifications), the taxonomic profiles associated with depth of the archaeal *narG* gene potentially warrants presenting here. At 1 cm, the *Euryarchaeota* accounted for 59% of the archaeal *narG* abundance including the *Thermoprotei* (32%) and the *Thaumarchaeota* (2%). Of the *Euryarchaeota* at this same shallow depth, the family *Archaeoglobaceae*, and the *Methanosarcinales* genera *Methanoperedens* and *Methanosaeta* accounted for 18%, 10%, and 10% of archaeal derived *narG* gene sequences respectively. Of the *Thermoprotei*, the family *Sulfolobaceae* account for 21% of the archaeal derived *narG*. At 31 cm, this pattern changed as the *Euryarchaeota* derived abundance increased to 85% of archaeal contribution, and the *Thermoprotei* and Thaumarchaeota *decreased* to 10% and 0.7% respectively. Of the *Euryarchaeota*, *Methanoperedens* contributed 34% of archaeal *narG* abundance, *Methanocella* 14%, and *Methanosaeta* 9%. Here the *Sulfolobaceae* (*Thermoprotei*) decreased to 4% of archaeal contribution.

3.3.8.3 *NiFe hydrogenase hydB GraftM Results*

The relative abundance of *hydB* gene sequence copies identified by GraftM increased as a proportion of the total sequences in each sample with depth from 0.042% at 1 cm to 0.11% at 31 cm. The NiFe hydrogenases are comprised of five main groups according to their role, composition, and localisation (Schäfer *et al.*, 2013) and these main groups are subdivided by a range of properties, including taxonomic and functional differentiation (Søndergaard *et al.*, 2016). The Group 4d peaked at 19.54% at 3 cm and decreased to 14.39% at 31 cm. An unspecified Group 4 group increased from 19.08% at 1 cm to 20.16% at 9 cm, and decreased to 11.35% at 31 cm. The Group 4e decreased from 14.95% at 1 cm to 8.5% at 31 cm. The Group 3b increased from 7.63% at 1 cm to 13.29% at 17 cm, and decreased to 11.28% at 31 cm. The Group 4f decreased from 12.02% at 1 cm to 2.71% at 31 cm. The Group 3c increased from 1.53% at 1 cm to 19.14% at 31 cm. The Group 3d increased marginally from 4.96% at 1 cm to 7.28% at 31 cm. An unlabelled group decreased from 3.88% at 1 cm to 1.97% at 31 cm. Group 1a increased from 0.95% at 1 cm to 8% at 31 cm. Group 1f increased from 1.21% at 1 cm to 3.78% at 17 cm and decreased to 3.21% at 31 cm. (Appendix C.3)

3.3.8.4 *Acetyl CoA synthase acsB GraftM Results*

The relative abundance of *acsB* gene sequence copies identified by GraftM increased as a proportion of the total sequences in each sample with depth from 0.001% at 1 cm to 0.027% at 31 cm. A bacterial *acsB* group increased 34.38% at 1 cm to 64.29% at 9 cm, and decreased to 43.9% at 31 cm. A δ-*proteobacteria acsB* group decreased from 28.13% at 1 cm to 6.11% at 17 cm and increased to 16.49% at 31 cm. An archaeal *acsB* increased from 3.13% at 1 cm to 13.08% at 31 cm. A *Nitrospira acsB* was 9.38% at 1 cm, 1.21% at 3 cm, increased to 10.03% at 23 cm, and decreased to 8.72% at 31 cm. A *Methanomicrobiales acsB* decreased from 6.25% at 1 cm to 0% at 9 and 17 cm, and increased to 5.45% at 31 cm. A *Chloroflexi acsB* decreased from 3.13% at 1 cm to 0.67% at 23 cm and increased to 2.73% at 31 cm. A *Methanomicrobia acsB* was recorded at 3, 17, and 31 cm (1.21%, 2.29%, and 2.04% respectively). A *Methanosarcinales acsB* decreased from 3.13% at 1 cm to 0% at 9 and 17 cm, and increased to 0.55% at 31 cm. An unlabelled *acsB* was 0% at 1 cm and 3 cm, 2.38% at 9 cm, 0% at 17 cm, and increased to 1.09% at 31 cm (Appendix C.4).

3.3.8.5 *Adenylylsulfate reductase aprA GraftM Results*

The relative abundance of *aprA* gene sequence copies identified by GraftM increased as a proportion of the total sequences in each sample with depth from 0.11% at 1 cm to 0.17% at 31 cm. The *Bacteroidia* WCHB1-69 decreased from 10.52% at 1 cm to 4.82% at 31 cm. An unspecific *Proteobacteria* increased from 5.41% at 1 cm to 7.15% at 17 cm and decreased to 1.93% at 31 cm. The *Firmicutes* had a relatively consistent abundance throughout, averaging 3.08 ± 0.1% across all six samples. A β-*proteobacteria* had an average abundance of 2.85 ± 0.14% between 1 – 23 cm which decreased to 1.68% at 31 cm. The *Burkholderiales* genus *Parasutterella* had an average abundance of 2.85 ± 0.22% between 1 – 23 cm, which decreased to 1.09% at 31 cm. The *Spirochaetes* genus *Treponema* had an average abundance of 2.01 ± 0.13% across all samples. The δ-*proteobacteria* genus *Syntrophorhabdaceae* increased from 0.04% at 1 cm, to 2.33% at 23 cm, and to 8.39% at 31 cm. The genus *Nitrospira* increased from 1.29% at 1 cm to 3.06% at 9 cm and decreased to 1.46% at 23 cm, and to 0.13% at 31 cm. An α-*proteobacteria* increased from 1.78% at 1 cm to 2.61% at 9 cm and decreased to 0.97% at 31 cm. A γ-*proteobacteria* increased from 0.61% at 1 cm to 2.33% at 23 cm and decreased to 1.26% at 31 cm (Appendix C.5).

3.3.8.6 *Methyl Coenzyme M Reductase mcrA GraftM Results*

The *mcrA* gene had very few sequence copies identified in the metagenome, with GraftM identifying as few as four copies of the *mcrA* at 1 cm, and only 51 copies at 31 cm. Only 16 taxa were identified in total, and all are presented in Appendix C.6. Most of these were zero counts and the remainder were single or few sequence copies from 1 – 23 cm. At 31 cm, the most abundant taxon was unspecified *Methanomicrobiaceae* (37.5%) followed by the *Methanomicrobiaceae* genus *Methanoregula* (29.17%), and then the *Methanosaeta* (10.42%) (Appendix C.6).

3.4 Discussion

A range of DNA-based molecular methods were used to characterise and construct depth-related profiles of the microbial community taxonomic and functional assemblages of the Needle's Eye contaminated and control cores. Based on these taxonomic and functional profiles of the cores, descriptions and inferences of the prevailing biogeochemical and redox conditions of the cores were made, and these were compared with directly observed and measured geochemical conditions of the same cores where possible. This included a focus on characterising the transport and redox fate of the uranium (U) in the contaminated core.

The microbial methods used, ranging from relatively rudimentary quantification of DNA and 'end point' PCR through to contemporary high-throughput DNA sequencing and sequencing analysis methods, were generally corroboratory in the profiles of microbial taxonomic and functional assemblages as a function of depth generated. Functional gene profiles generated by the range of methods appear to correlate with one another in the contaminated core, while there is less correlation in the control core, likely due to the different 16S rRNA marker gene sequencing method used for that core as well as the lack of shotgun metagenome sequencing analysis. That abundance profiles have consistent, progressive, and fine-scaled patterns with increasing depth reproducible across the methods which enhances the confidence that these community compositions are not merely random in their distributions throughout the cores and are indeed structured based on growth and selection that reflects the stability and likely impact of and on the surrounding geophysical, hydrological, and geochemical environments.

Within the Needle's Eye site, the two cores appeared to have similar overall depth-related stratification of microbial taxonomic and functional abundance and indeed contained many of the same taxonomic groups and functional genes. Both cores appeared to show evidence of the normally-expected, diffusion and thermodynamically-controlled, vertical stratification of redox function in soils and sediments (Thauer *et al.*, 1977). Overall, the control core appeared to have relatively gradual succession with depth, while the contaminated core showed more of a step-change at ~ 12 cm in its profile changes. Ordination plots of the 16S rRNA marker gene sequencing analysis appeared to consistently show progressive and gradual change in community structure with depth in both cores from surface to ~ 20 cm, after which community structure appeared to become more homogeneous. The

microbial taxonomic and functional profiles indicate that the microbial communities are probably dictated and controlled by the flow of subsurface oxic or microoxic groundwaters at the site (MacKenzie *et al.*, 1991; Jamet *et al.*, 1993). Additionally, carbon, specifically the balance between refractory and labile carbon availability in the site, likely has a role in microbial community structure and function and will be discussed throughout as a function of each core region.

The control core was found to contain significantly lower levels of U compared to that of the contaminated core, but a lack of equivalent geochemical analysis and shotgun metagenome sequencing of the core means that definitive conclusions and corroboration between microbial and geochemical profiles was not possible. The control core did however, due to the apparent relative similarity in community and functional composition between the control and contaminated cores, imply that there was no major impact of metal toxicity arising from the relatively high levels of U and Zn in the contaminated core.

3.4.1 The Contaminated Depth Core

The contaminated core will be discussed in terms of three clearly-defined depth sections each characterised by marked changes in the dominant microbial taxonomic and functional abundances that emerged from molecular analysis. Briefly, they are: The upper 1 – 23 cm region, as characterised by a transition with depth from initially mixed function, to aerobic, to microoxic conditions; 24 – 34 cm region, which represented a relatively spatially-short transition to highly reducing conditions with the presence of methanogenic and sulfate-reducing taxa, and; the 35 cm interval which was, interestingly, characterised by a highly-localised emergence of anaerobic methane-oxidising and other microaerophilic taxa and functions.

Overall patterns of the contaminated core show a gradual and progressive decrease in DNA and TOC levels with depth (Figures 3.1A and 3.1B). The correlation of these DNA and TOC values (Figure 3.2) could potentially be evidence of higher carbon availability directly leading to higher microbial activity and growth and therefore higher diversity (Drenovsky *et al.*, 2004). Alpha diversity index values were also highest in the surface samples and all decreased with depth (Figure 3.6), indicating a pattern of decreasing species richness, diversity, and evenness with increasing depth. This may potentially be linked to the DNA and TOC levels,

indicating that carbon (C) becomes more of a limiting factor at depth, and limits microbial community diversity. Environmental fitting of TOC and metal values appeared to consistently show correlation between TOC, U, Zn, and Mn levels with the upper 1 – 23 cm samples (Figures 3.27 and 3.28). These data indicate the composition, distribution and abundance of organic matter is a major driver of biogeochemical process.

The range of molecular methods were used in combination to draw conclusions and inferences of the prevailing geochemical conditions of the core, rather than the profiles generated from each method being directly compared with one another.

3.4.1.1 *The 1 – 23 cm Region of the Contaminated Depth Core: Putatively Aerobic*

Microbial community assemblage of the 1 – 23 cm region of the contaminated core showed the *Acidobacteria* subgroup 6 to be the most abundant taxonomic group, peaking at 9.39% of the microbial community assemblage at 5 cm (Figure 3.14). *Acidobacteria* are a poorly-characterised phylum with no cultured isolate (Hug *et al*., 2016) that are widespread and abundant in a range of soil environments (Kielak *et al.*, 2016), with the subgroup 6 in particular being associated with uranium-impacted soils (Barns *et al.*, 2007) and elevated Ca, Mg, Mn, and B soil content (Navarrete *et al.*, 2012). Though this was a highly abundant taxon in both cores, it does not help inform the geochemical processes of the site, as it is an insufficiently sequenced, cultured, or characterised taxonomic group to allow useful functional inferences. This same group was also found in the Control Core in high abundance, implying it is not selected by U.

The initial surface 1 – 5 cm zone of the contaminated depth core appeared to contain taxa and genes indicative of a mix of respiratory function. ‘End point’ PCR appeared to show abundance of denitrifying *nirS* and *nirK*, the oxygen-dependent bacterial *amoA* (albeit low band intensities) and archaeal *amoA* increasing in this zone, and, contradictorily, the presence *mcrA* gene in this zone usually associated with highly reducing conditions (Figure 3.4). Consistently, 16S rRNA marker gene sequencing taxonomic profiles appeared to show increasing abundances of the ammonia oxidising *Thaumarchaeota* family *Nitrosopumilaceae* (3.14% at 10 cm, Figure 3.15A) (Qin *et al.*, 2017), nitrite oxidising *Nitrosomonadaceae* (1.87% at 16 cm, Figure 3.15C) (Prosser, 2014), sulfur oxidising *Sulfurifustis* (4.55% at 20 cm,

Figure 3.14) (Kojima *et al.*, 2015) in this zone, as well as the presence of the methanogenic *Euryarchaeota* genus *Methanoregula* (0.64% at 3 cm, Figure 3.15F), corroborating the observed *mcrA* abundance in this region. An uncultured 'methanogenic' (as per the taxonomy string assigned during analysis) *Bathyarchaeia* was present in this zone, and more broadly throughout the 1 – 23 cm region (2.11% at 2 cm, Figure 3.14). The *Bathyarchaeia* were shown to contain *mcrA*, but their specific role in methane cycling is not yet clear (Evans *et al.*, 2015). Also, corroboratively, Tax4Fun showed the *amoA* gene increasing from 1 – 4 cm (0.96%), the *nirK* peak at 2 cm (4.14%) and a relatively high abundance of *narG* (7.65% at 1 cm) and *nosZ* (2.66 ± 0.07% between 1 – 23 cm), as well as a small peak in the *mcrA* gene between 1 – 5 cm (1.68% at 2 cm) (Figures 3.18 and 3.20). This coexistence of aerobic, oxidising and anaerobic, reducing taxa and genes likely reflected the diffusion of oxygen from the surface and its rapid consumption due to the high carbon levels, rapid depletion of oxygen and other terminal electron acceptors (TEAs) through to methanogenesis in small, highly-localised environments. The lack of evidence for sulfate reduction but the potential for methanogenesis might reflect low sulfate levels common in terrestrial habitats or the potential presence of methanogens associated with single or multicellular *Euryarchaeotes* (Bogard *et al.*, 2014).

By 8 cm, the ammonia oxidising taxa and genes had peaked, the nitrite and sulfur oxidising taxa continued to increase, denitrifying genes persisted, and the methanogenic taxa and genes had receded. This apparent transition from mixed aerobic and anaerobic processes between 1 – 5 cm to apparently exclusively aerobic processes at 8 cm is perhaps evidence of the subsurface horizontal groundwater flowing into the bog at this depth (MacKenzie *et al.*, 1991; Jamet *et al.*, 1993). This water is likely oxygenated, and hence sustains the presence of oxidising taxa and genes.

From 9 cm to 23 cm, a gradual decline in ammonia, nitrite, and sulfur oxidising taxa and genes occurred. Furthermore, putatively fermentative *Anaerolineaceae* (Yamada *et al.*, 2006) continued to increase with depth, and by 12 cm, δ-*proteobacterial* taxa and the methanogenic *Euryarchaeota Methanolinea* and *Bathyarchaeia* emerged. Denitrification genes appeared to persist until 23 cm. This transition from strictly-aerobic processes to the emergence of anaerobic processes is evidence of the normally-expected vertical stratification of redox processes with

depth – as oxygen that has diffused from 8 cm is consumed at depth, denitrification continues, and the removal of oxygen with depth enables the emergence of the strictly anaerobic taxa and genes observed.

On the basis of the previous discussion it would be expected from the taxonomic and functional profiles derived from this region that the 1 – 5 cm depth, due to the presence of strictly anaerobic taxa and genes, show some levels of microbially-mediated NO_3^{2-}, Fe^{3+}, and U^{6+} reduction, but that due to the oxidising conditions and likely subsurface groundwater flow at 8 cm, these reduced species would be rapidly re-oxidised at 8 cm. From 9 – 23 cm, the progressive decline of aerobic, oxidising taxa and genes in lieu of anaerobic, reducing taxa and genes, would likely result in the gradual reduction of TEAs as far as Fe^{3+} and U^{6+} reduction.

Geochemical analysis reflected these assumptions made for biogeochemical cycling in this region made from the microbial taxonomic and functional profiles. Bulk Mn and Fe profiles (Figures 3.26B and 3.26C) showed the loss of solid-phase Mn occurring progressively with depth immediately from the surface, and solid-phase Fe seeming to increase slightly initially from 1 – 8 cm, before declining with depth after ~ 10 cm. Critically, the $Fe^{3+/2+}$ ratio (Figure 3.26E) appeared to show all the Fe in the core was in the reduced Fe^{2+} form between 1 – 5 cm, before the emergence of oxidised Fe^{3+} which increased in percentage from 6 – 15 cm, before reduced Fe^{2+} quickly dominating again below this level. It is the coincidence of aerobic microbial groups in the zone of increased ferric iron that indicates that this anomalous result which does not conform to the typical vertical transition to progressively more reducing conditions is a genuine reflection of the system. Of particular interest and again entirely consistent with the microbiological patterning, U levels peaked at 8 cm, with the proportion of U^{6+} compared to reduced U^{4+} increasing to 92%, with the oxidised fraction persisting with depth, to 80% at 23 cm (Table 3.4). The mechanism by which this oxidised, thus potentially most soluble and hence mobile form of U, is retained in the most oxic part of the bog, rather than being transported in the groundwater to ultimately be transported to the nearby estuary, is thought to be related to complexation with indigenous refractory organic matter at the site and the exact mechanism which involves the formation of carbonyl linkages will not be explored further in this thesis (Fuller *et al.*, in prep.). However, a number of important features of the U retaining layer within this core can be inferred from the specific hydrology, microbiology and carbon composition/distribution in the Needle's eye U

retaining bog soil. For instance, it is reasonable to infer relatively long-term constant oxidising conditions at the 10 cm depth increment as evidenced by the distinctive microbial community which has been selected at this depth. This condition is likely sustained by the horizontal flow of oxic waters carrying U^{6+} into the soil, however, for this to be feasible the abundant organic matter is most likely refractory as otherwise oxygen would most likely have been consumed before reaching the core location. Thermograviometric and differential scanning calorimetry (TG-DSC) analysis carried out on the bulk upper layer of the soil of the contaminated and control cores (Clare McCann pers. com., Appendix B3) showed that most of the energy released was at temperatures above 400˚C, implying that the bulk of the organic carbon in these soils is recalcitrant, and likely comprised of lignin and polyphenols rather than the lower temperature combustible components such as cellulose (Lopez-Capel *et al*., 2006; Blake *et al*., 2017). This dominance of recalcitrant organic carbon is in fact likely the consequence of this sustained horizontal oxic flow itself whereby that preservation of organic matter in the surface layer is by stabilisation through aerobic composting by the action of aerobic bacteria as evidenced by the peaking of the inferred presence of cytochrome oxidases at this depth and the dominance of lithoautrophic metabolism, *e.g.*, ammonia and sulfur oxidation. This situation contrasts with the previously assumed mechanism for U retention at this site by organic matter preservation and reduced uranium retention resulting from anaerobic conditions (MacKenzie *et al.*, 1991).

3.4.1.2 *The 24 – 34 cm Region of the Contaminated Depth Core: Putatively Anaerobic*

All microbial methods consistently identified patterns in the microbial community's taxonomic and functional assemblages changing in a very small spatial-scale between the 1 – 23 cm and 24 – 34 cm regions. This change was characterised by an almost-complete removal of aerobic, oxidising taxa and associated genes such as the *Thaumarchaeota* and *Acidobacteria*, and the significant increase in sulfate-reducing and methanogenic taxa and associated genes. This was reflected in the 'end point' PCR, 16S rRNA marker gene sequencing, Tax4Fun profiles, and shotgun metagenome sequencing analyses.

The microbial community assemblage became more homogeneous compared to the 1 – 23 cm as per the ordination plots, less species rich, diverse, and even as

assessed by alpha diversity metrics. DNA and TOC levels also dropped significantly in this region compared to the 1 – 23 cm region also.

As per Figures 3.14 and 3.15, the most abundant taxonomic groups in this region were the δ–*proteobacteria* Sva0485, a potential sulfate-reducer (Concheri *et al.*, 2017), the sulfate-reducing *Thermodesulfovibriona* (Frank *et al.*, 2016), and archaea including the methane cycling *Bathyarchaeia* (Evans *et al.*, 2015) order *Thermoprotei* and the hydrogenotrophic methanogenic *Methanolinea* (Imachi *et al.*, 2008; Sakai *et al.*, 2012). Interestingly, the *Thermodesulfovibriona* is an incomplete oxidising SRB, producing acetate but can also use H_2 as an electron donor (Sekiguchi *et al.*, 2008). Furthermore, *Methanolinea* which uses H_2 also were higher in abundance than the acetoclastic *Methanosaeta* (Ma *et al.*, 2006) and *Desulfobacca* (Oude *et al.*, 1999), indicating the potential for chemolithoautotrophic processes in this region of the core, similar to the 1 – 23 cm region, despite the still relatively high TOC values in this region.

Functional gene abundance in this region appeared to transition from oxidising and denitrifying genes to sulfate-reducing and methanogenic genes, across 'end point' PCR, Tax4Fun, and shotgun metagenome sequencing. Genes from the dissimilatory sulfate reduction pathway *aprA* and *dsrA*, methanogenic *hdrA* and *mcrA*, and anaerobic carbon-fixating *cooS* and *acsB* all significantly increased in this region. GraftM showed a transition to *Desulfobacteraceae narG* taxonomy. The denitrifying *narG*, *nirK*, and *nosZ* all significantly declined while the dissimilatory nitrate reduction (DNRA) *nrfA* and nitrogen-fixation *nifD* increased. This is interesting because *nrfA* requires nitrite the same as *nirS* and *nirK*, so the presence and abundance of this gene at depth implies nitrite is still available as a TEA, while the abundance of *nifD* implies the fixation of N_2, the gaseous end product of denitrification. Both DNRA and nitrogen fixation produce ammonium as an end product, which can be more readily oxidised into nitrite than N_2 (Gray and Head, 2005). This switch from denitrification at 1 – 23 cm to DNRA and nitrogen fixation between 24 – 34 cm likely reflects the imposed anaerobic conditions of the 24 – 34 cm region, with lower levels of nitrate. This is DNRA is no less thermodynamically favourable when compared to denitrification when the system is limited with respect to this electron acceptor (Gray and Head, 2005).

The microbial taxonomic and functional profiles leave little doubt to the biogeochemical functions occurring in this region of the core, which is primarily

sulfate reduction and methanogenesis, as well as DNRA and nitrogen fixation. This region of the core is clearly very reducing, and all more favourable TEAs from oxygen to Fe^{3+} and U^{6+} would be expected to have been exhausted. Low sulfate levels common in terrestrial environments may explain the relatively high abundance of methanogenic taxa and genes when sulfate reduction is generally more thermodynamically favourable than methanogenesis and would normally outcompete methanogens for available electron donor.

Geochemical analysis confirmed the assumptions informed by the microbial data. Solid phase Mn is absent, indicating a complete reduction of Mn^{4+} into soluble Mn^{2+} (Figure 3.26B). Solid phase Fe declined after 20 cm (Figure 3.26C), while $Fe^{3+/2+}$ ratio showed the total Fe stock in this region to be very nearly 100% reduced Fe^{2+} (Figure 3.26E). Porewater U levels are significantly lower in this region (Figure 3.26D, Appendix A.14A) presumably because of the absence inflowing oxic U contaminated waters. However, of the U detected between 70 – 90% of the U at 30 cm was reduced U^{4+} (Table 3.4) consistent with the prevailing conditions being dominated by sulfate reduction and methanogenesis. The mechanism by which up to 30% of the U persists as the oxidised U^{6+} in an otherwise putatively reducing region of the core is thought to again be evidence of the refractory nature of the organic matter, and the complexation of the U with it (Fuller *et al.*, in prep.).

The relatively short spatial transition in biogeochemical conditions observed around 23 – 24 cm may reflect and be imposed by the subsurface hydrological properties of the site as described by MacKenzie *et al.* (1991) and Jamet *et al.* (1993). If subsurface oxygenated groundwater was flowing horizontally at ~ 8 cm depth, and the diffusion of oxygen downwards was occurring as is evidenced by the presence of oxidising taxa and genes, this will inhibit O_2-sensitive processes such as methanogenesis. However, as soon as O_2 and other relatively oxidising TEAs such as nitrate are depleted, geochemical conditions become more reducing. The emergence of sulfate-reducing and methanogenic processes at 24 cm reflect a depletion of O_2, and the very short spatial transition zone may reflect sufficient labile C content for the reduction of more favourable TEAs including nitrate, and Fe as measured by XRF. Higher resolution sampling of the soil in this region, *e.g.* in the 0.5 cm scale, might have elucidated a more progressive and specific stratification in this transition zone.

One interesting observation in this deeper region of the core was the increase in the NiFe hydrogenase *hydB* gene in this 24 – 34 cm region, particularly in the group classification of these gene fragments as detected by GraftM. NiFe hydrogenases are a part of the microbial H_2 cycle that reversibly cleave H_2 into two protons and two electrons, enabling H_2 to act as an electron donor for respiration or fermentation (Schäfer *et al.*, 2013). Between the 1 – 23 cm and the 24 – 34 cm region of the contaminated core, there was both an increase in NiFe hydrogenase abundance (Figure 3.21E, Appendix A.10E) and a general transition from Group 4 (Group 4d) to Group 3 (Group 3c in particular) was observed with increased depth (Appendix C.3). The Group 4 NiFe hydrogenases are mostly associated with archaea including *Methanosarcina* and produce H_2 from the reduction of protons and oxidation of C_1 compounds in anaerobic environments (Vignais and Billoud, 2007). While the *Methanosarcina* were not presented in the 16S rRNA gene sequencing taxonomic profiles of the 24 – 34 cm region of the core, they were present between 1 – 23 cm, ranging between 0.68% at 3 cm, to 0.16% at 16 cm, 1.02% at 21 cm, but 0% by 29 cm. The Group 3 NiFe hydrogenases meanwhile are associated with archaea more generally (Vignais and Billoud, 2007) and Group 3c is involved in the bifurcation of electrons from H_2 to heterodisulfide and ferredoxin, the final step in the Wolfe cycle of hydrogenotrophic methanogenesis (Thauer, 2012). This transition to hydrogenotrophic *hydB* NiFe hydrogenase suggests the presence of hydrogenotrophic methanogenic processes in the 24 – 34 cm region as discussed in the previous section (3.4.1.2.).

3.4.1.3 *The 35 cm Region of the Contaminated Core: Anaerobic Methane Oxidation*

Subsurface groundwater flow at the Needle's Eye site has been described by MacKenzie *et al.* (1991) as occurring at multiple depths in the site. At 35 cm in the contaminated core a consistent pattern, characterised by the sudden decline in anaerobic and reducing taxa and genes with a simultaneous emergence in a range of aerobic and oxidising genes and taxa similar to the 1 – 23 cm region, was identified by all the methods used. This is likely to be evidence of sustained (based on clear community selection and microbial) oxygenated, or at least microaerophilic groundwater flow at this particular depth increment. The specific biogeochemical processes and nature of this increment are found to be intriguingly distinct from oxidised upper layers and will be explored further.

The microbial community assemblage at this 35 cm increment was identified by beta diversity ordination to be less similar to the 24 – 34 and 36 – 41 cm samples, but also not similar enough to the 1 – 23 cm region to be plotted amongst those samples. The 35 cm depth increment therefore appears to be an outlier in terms of its community assemblage (Figures 3.10 and 3.11). More specifically, a range of taxa and genes emerged almost exclusively at this depth. Most interestingly the methylotrophic *Methylomirabilis* (Figure 3.15B) and *Methylocystis* (not presented in the taxonomic analysis) genera were almost exclusively abundant at 35 cm. Genes including the *mxaF* methanol dehydrogenase (Figure 3.21D), and the denitrifying *napA* (Figure 3.20B), *nirS* (Figure 3.19B), and *norB* (Figure 3.20C) genes were relatively enriched almost exclusively at this depth according to Tax4Fun. Other genes, including the *mmoX* methane monooxygenase (Figure 3.21C), ammonia oxidising *amoA* (Figure 3.18A), and denitrifying *narG* (Figure 3.20A) and *nirK* (Figure 3.19C) also re-emerged at this depth increment. All anaerobic genes described as abundant in the 24 – 34 cm region, including the *dsrA* (Figure 3.18D) and *mcrA* (Figure 3.18E), were largely absent at 35 cm. These trends in the Tax4Fun profiles were broadly reflected in the 'end point' PCRs also (Figure 3.4, Figure 3.19). Shotgun metagenome sequencing was not performed on the 35 cm sample, but at 31 cm, GraftM analysis showed 2% of the *narG* gene sequences classified as belonging to *Methanoperedens*. However, *Methanoperedens* was not detected in the 16S rRNA marker gene sequencing analysis, presenting a conflict in the taxonomic profiles.

Methylomirabilis is an anaerobic methane-oxidising bacterial genus that couples anaerobic methane oxidation with nitrite reduction, via the production and use O_2 from the disproportionation of nitrite (Wu *et al.*, 2011). However, it has been shown that exposure of *Methylomirabilis* to high levels of oxygen induces an oxidative stress response from the genus (Luesken *et al.*, 2012). Intriguingly, the genus *Methylocystis* is a bacterial aerobic methane oxidiser which uses O_2 to oxidise CH_4 and outcompetes heterotrophs for carbon as an electron donor at low carbon levels (van Bodegom *et al.*, 2001). *Methanoperedens* is an archaeal genus, previously described as ANME-2d, that performs anaerobic methane oxidation via a reverse methanogenesis pathway, using nitrate as the terminal electron acceptor, and is commonly associated with the *Methylomirabilis* taxa in anoxic soils (Haroon *et al.*, 2013, Vaksmaa *et al.*, 2017).

The genes *mmoX* and *amoA* both are associated with the first step of the methane oxidation pathway of oxidising methane (CH_4) into methanol (CH_3OH), while *mxaF* performs the second step of methane oxidation by oxidising methanol into formaldehyde, the terminal product in methanotrophic methane oxidation. The denitrifying pathway genes *napA*, *nirS*, and *norB* were all enriched at 35 cm, but not the terminal denitrifying enzyme *nosZ* which catalyses nitrous oxide to N_2. The absence of *nosZ*, but the presence of these other denitrifying genes, is consistent with metagenome studies of *Methylomirabilis*, where *M. oxyfera* was shown to contain all the above genes except *nosZ*, and instead is thought to perform nitrous oxide to N_2 reduction by converting two nitric oxide molecules directly into N_2 and oxygen (Figure 3.30, from Ettwig *et al.*, 2010). The relationship between *Methanoperedens* and *Methylomirabilis* is interesting because it appears that *Methanoperedens* reduces nitrate to nitrite and N_2 (releasing O_2) and *Methylomirabilis* reduces nitrite to N_2, both using CH_4 as their electron donors, and both producing CO_2 as a by-product (Vaksmaa *et al.*, 2017; Ettwig *et al.*, 2010). The release of O_2 from anaerobic methane oxidation by *Methylomirabilis* (and *Methanoperedens* as inferred from the *narG* phylogeny) as well as any residual O_2 in the groundwater may explain the enrichment of the aerobic methane-oxidising bacteria *Methylocystis* at 35 cm.

Based on the microbial taxonomic and functional assemblage at 35 cm described above, the presence of both aerobic and anaerobic methane oxidation pathways is inferred. The following equation from Ettwig *et al.* (2010) shows the thermodynamic equation of the reaction of methane with nitrite (and nitrate), while Figure 3.37 shows the significant pathways of *Methylomirabilis oxyfera*.

$$3CH_4 + 8NO_2^- + 8H^+ = 3CO_2 + 4N_2 + 10H_2O$$

$$(\Delta G^{o\prime} = -928\ kJ\ mol^{-1}\ CH_4)$$

This highly localised enrichment of a range of oxidising and denitrifying taxa and genes at 35 cm, in a region of the core that is generally dominated by anaerobic and reducing function is unusual and is likely further evidence of subsurface groundwater flow introducing fully oxic or, more likely based on the anaerobic methane oxidation taxa present, microaerobic, conditions at this depth increment. Furthermore, there was evidence of U reoxidation at this depth, though the two methods used disagree on how much, with L_3-edge showing 27% of the U as oxidised U^{6+}, while M_4 showed as much as 75% (Table 3.4).

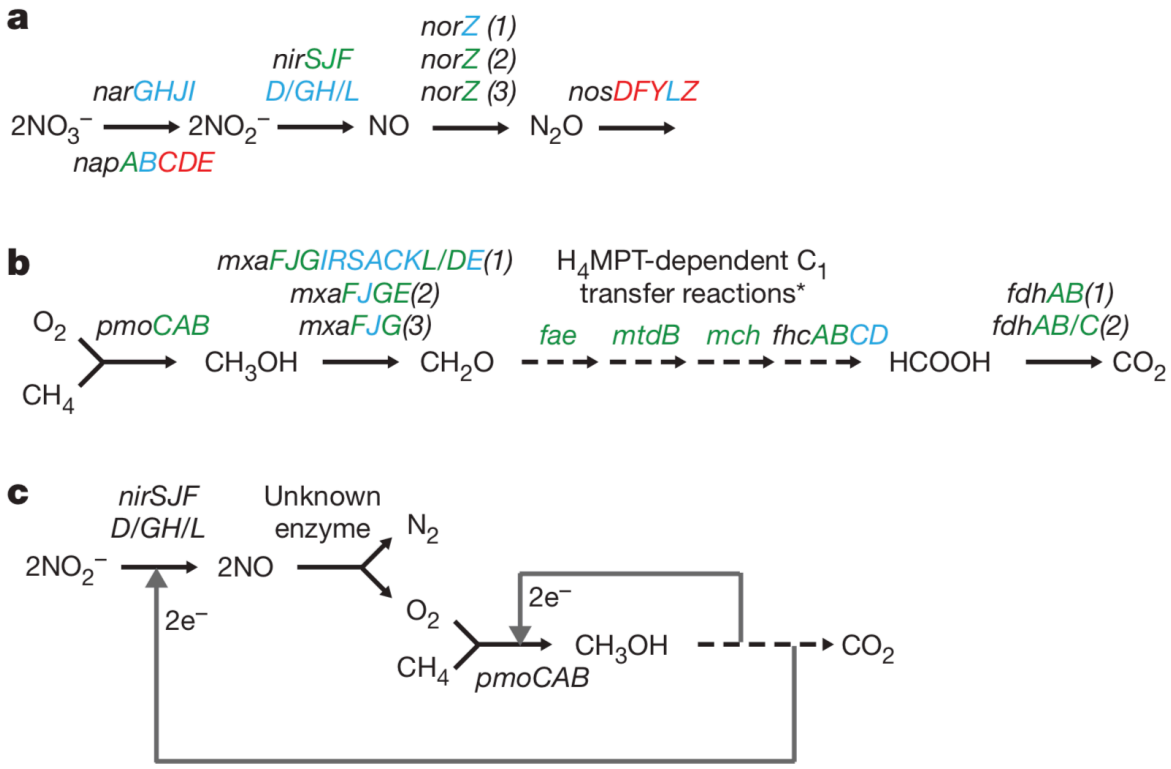


Figure 3.30. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. From Ettwig *et al.* (2010).

3.4.2 The Control Depth Core

Although there was a lack of geochemical analysis performed on this control depth core similar to that done for the contaminated core, it was discovered from bulk analysis that TOC content in this section of the Needle's Eye site was ~ 60% (Adam Fuller, Pers. Com.), while U levels were not found to exceed 10 ppm – much lower than the ~ 1600 ppm peak observed in the contaminated core, making the control core a useful control from a U-contamination perspective. However, in hindsight, the likelihood of finding a depth core with identical geophysical, geochemical, and hydrological characteristics to the contaminated core but without U was always going to prove difficult, even in the Needle's Eye site. The large disparity in the range of different methods used between the two cores also hindered a truly fair comparison. Despite these shortcomings in the experiment, the control core proved to be an effective counterpart to the contaminated core, and similarities and dissimilarities in the microbial community assemblages and inferred functions will be discussed.

The control depth core appeared to share common characteristics with the contaminated core, including similar alpha diversity index values which declined with

depth (Figures 3.6 and 3.7), a progressive distribution of microbial community assemblages associated with increasing depth (Figures 3.8 and 3.9), abundant taxa, *e.g.*, the *Acidobacteria* subgroup 6, *Anaerolineaceae*, *Bathyarchaeia*, and *Thermodesulfovibriona* (Figure 3.15), and similar profiles of functional gene abundances (Figures 3.4, 3.22, 3.23, and 3.24). However, notable differences between the structure and depth-related distribution patterns emerged between the two cores. Where the contaminated core had very clearly defined regions of putative microbial function and biogeochemical cycling that generally followed the normally expected vertical thermodynamically controlled stratification of redox processes, the control core appeared to have a more gradual and progressive transition in microbial community assemblage that was less definitive. Ordination reflected this more gradual transition, generally characterised in order of the upper 1 – 11 cm and 12 – 18 cm regions, with the 20 – 41 cm being more homogenous similar to the 24 – 41 cm (excluding 35 cm) region in the contaminated core (Figures 3.12 and 3.13). Furthermore, putative microbial function at depth in the control core appeared to be a more heterogeneous mix of aerobic and anaerobic processes.

However, despite the differences in microbial community stratification with depth between the two cores, it is likely that the hydrology of the Needle's Eye site is still a determining factor in this stratification, and that the hydrology and subsurface oxic groundwater flows are less pronounced or obvious in this control core. Although carbon levels were lower in the control core than the contaminated core (60% in the control and as high as 80% in the contaminated core), the relatively higher abundance of chemoheterotrophs throughout the control core compared to the higher abundances of chemoautotrophs in the contaminated core suggest that the carbon in the control core is potentially more labile.

This discrepancy between 'end point' PCR band intensities and Tax4Fun inferred abundance profiles for the *amoA* and *mcrA* (Figure 3.23) may be due to two reasons. Firstly, and more pertinent to the *mcrA* specifically, it may be due to the lack of *Euryarchaeota* in the 16S rRNA gene taxonomic profiles. Tax4Fun uses precomputed matrices of 16S rRNA taxonomy and KEGG function and uses these matrices to infer functional profiles. If a taxonomy linked to a particular function is not present in the input OTU table, that function will not be observed. Conversely, as databases are updated and new function is linked to taxonomy, but the Tax4Fun matrices are not updated, the Tax4Fun profiles will not reflect this new function.

Archaea are observed in the control core, but only the *Bathyarchaeia* and *Thaumarchaeota*, whose function has been elucidated more recently (Evans *et al.*, 2015; Pester *et al.*, 2011). While the *Bathyarchaeia* are thought to contain a version of the *mcrA* gene, this must not be reflected in the Tax4Fun matrices. Secondly and potentially relatedly, the control core was sequenced on the Ion Torrent platform targeting the longer V4-V5 region of the 16S rRNA, generating longer single-end reads of much varying length and quality compared to that generated from Illumina sequencing for the contaminated core. Illumina sequencing targeted the V4 region of the 16S, and produced shorter, paired-end reads of a specific length. It is possible that the Ion Torrent sequencing does not target archaeal 16S rRNA sequences as well as the Illumina, either due to the variable length or greater range of quality of sequences.

3.4.2.1 *The 1 – 11 cm Region of the Control Core*

Similar to the contaminated core, the most abundant taxonomic groups in this region of the control core included the *Acidobacteria* subgroup 6 (8.39% at 3 cm), the *Anaerolineaceae* (3.05% at 10 cm), and the NC10 *Rokubacteriales* (3.4% at 7 cm) (Figure 3.16), as well as the *Chloroflexi* KD4-96 (2.4% at 9 cm, not in the profile figures). The putative denitrifying δ-*proteobacteria* MBNT-15 (Hug *et al.*, 2016) is also present in this section, increasing to 3.33% at 10 cm.

Target functional genes within the control core showed mixed depth abundance profiles, with poor correlation between the ‘end point’ PCR gene target band intensities and Tax4Fun inferred relative abundances of the same genes, especially between *amoA* and *mcrA* genes (Figure 3.19). However, it is the ‘end point’ PCR profiles that are most similar between the contaminated and control core, and the profiles of the ‘end point’ PCR target genes seem to indicate vertical stratification of redox processes, with *amoA*, *nirS*, and *nirK* being most abundant in this 1 – 11 cm region, and the *dsrA* and *mcrA* relatively low in this region (Figure 3.5).The Tax4Fun profiles of the *amoA* gene appeared to be abundant in 1 – 10 cm region and decreased after 10 cm, but then gradually increased consistently with depth (Figure 3.22A), while the *mcrA* had no pattern with depth and consisted of many zero counts (Figure 3.22E). The *nirK* gene had a statistical correlation in abundances between ‘end point’ PCR and Tax4Fun, and the *dsrA* looked to visually correlate.

Other genes targeted in the Tax4Fun profile showed the denitrifying *narG* and *norB* genes abundant in this region, and with the *nirK* being abundant and the terminal denitrifying *nosZ* gene shows the presence of a complete denitrification pathway (Figure 3.24). Additionally, the oxygen-dependent *mmoX* and, unexpectedly, the methanogenic *mtdB* gene, were both more abundant in this region (Figure 3.25).

Based on the microbial community and functional profiles, it is inferred that the 1 – 11 cm region of the control core is aerobic and oxidising in function at the surface with the presence of denitrification, and the gradual increase in *dsrA* and *mcrA* indicating the gradual introduction of more reducing conditions at depth.

3.4.2.2 *The 12 – 18 cm Region of the Control Core*

Rather than the short spatial transition from denitrifying conditions to methanogenic and sulfate conditions observed in the contaminated core, the control core had a more gradual and progressive transition with depth between 12 – 18 cm (Figures 3.12 and 3.13). This region saw an increase in an uncultured methanogenic *Bathyarchaeia*, *Acidobacteriales*, the aerobic chemolithotrophic iron-oxidising *Gallionellaceae* (Hallbeck and Pedersen, 2014), and the sulfate-reducing *Thermodesulfovibriona* (Sekiguchi *et al.*, 2008) (Figure 3.16), in addition to the sulfate-reducing *Desulfobacca* (Oude *et al.*, 1999) and aerobic *Pseudolabrys* genera (Kämpfer *et al.*, 2006) (Figures 3.17A and 3.17B), with a decline in the *Acidobacteria* Subgroup 6 (Figure 3.16). ‘End point’ PCR showed the *amoA*, *nirK*, and *nirS* decline in this region, and the *dsrA* and *mcrA* increase (Figure 3. Tax4Fun showed a decrease in the same genes, and an increase in the *dsrA* (but not the *mcrA* as explained above) (Figure 3.22). Other denitrifying genes had a mixed profile, with the *narG* and *nosZ* decreasing but the *napA* and *norB* increasing (Figure 3.24A–D). The DNRA nitrite reductase *nrfA* and nitrogen-fixing *nifD* increased with depth as they did in the contaminated core (Figure 3.24E and F). Other anaerobic genes such as the *acsB*, *aprA*, *hdrA*, and *cooS* increased with depth (Figure 3.25).

The inferred function of this region of the core appears to be one of mixed aerobic and anaerobic processes. Though the oxygen-dependent *amoA* and *mmoX* genes and the denitrifying genes decreased in this region, the microaerobic iron-oxidising *Gallionellaceae* increased (Albuquerque *et al.*, 2011; Emerson, 2012). The *Acidobacteria* ASV is also likely to be aerobic, as per the *Acidobacterium capsulatum* (Kishimoto *et al.*, 1991). The *Gallionellaceae* and *Acidobacterium* and *Desulfobacca* Are chemoheterotrophic (Oude *et al.*, 1999). TOCs were found to be ~ 60% in the

control core. The higher abundances of chemoheterotrophs relative to chemoautotrophs in this core is the opposite to what was observed in the contaminated core and may indicate that the carbon in the control core is potentially more labile in nature in this region.

3.4.2.3 *The 19 – 41 cm Region of the Control Core*

This region of the core continues the trend observed in the 12 – 19 cm region of increased mix of aerobic and anaerobic taxa and function. The *Thaumarchaeota* ammonia-oxidising genus *Candidatus Nitrosotalea* (Prosser and Nichol, 2015), the *Gallionellaceae* iron-oxidising genus *Sideroxydans* (Weiss *et al.*, 2007; emerged with increased depth (Figure 3.17). The α-*proteobacteria* that was matched with a *Rhodospirallaceae* sequence in BLAST, methanogenic *Bathyarchaeia*, and the iron-oxidising *Gallionellaceae* (Hallbeck and Pedersen, 2014) dominate in this region. The obligately anaerobic fermenting *Micropepsaceae* isolated from a nutrient poor fen (Harbison *et al.*, 2017) was also abundant in this region (Figure 3.17C). The *amoA* profiles increased (Figures 3.5A, 3.22A) at the bottom of the core, potentially reflecting the increased *Nitrosotalea*. The *nirS* increased according to Tax4Fun (Figure 3.22B and 3.22D). The *mxaF* gene increased (Figure 3.25D) in this region, potentially reflecting the *Pseudolabrys* (Kämpfer *et al.*, 2006). Anaerobic genes including the *dsrA*, *aprA*, *acsB*, *hdrA*, and *cooS* continued to increase with depth.

This heterogeneous mix of strictly-aerobic and -anaerobic taxa and genes presents a conflict in accurately predicting the prevailing geochemical conditions of the core. It is possible that the core is aerobic and oxidising, perhaps due to a slow but continuous input of oxygenated groundwaters, and that stability in the system enables microenvironments of anaerobic function to exist.

3.4.3 The Fate of Radionuclides at Needle's Eye

Using a range of microbial tools, including 'end point' PCR and 16S rRNA gene sequencing, it was possible to accurately predict the redox speciation of U throughout the contaminated core. As discussed above, using the consistent, progressive, fine-scaled, and repeatable patterns of microbial taxonomic and functional distribution observed throughout this core, it was possible to deduce that this core was stratified into two main halves. The core transitioned relatively rapidly from aerobic and oxidising conditions in the upper ~ 0 – 22 cm region, to anaerobic

and reducing conditions between 23 – 41 cm, with an incursion of oxidising conditions localised to 35 cm. These profiles were used to make predictions of the prevailing geochemical conditions throughout the core, and critically, anticipate the speciation of U throughout the core. HERFD-XANES profiles produced by the University of Manchester corroborated these predictions accurately.

Using 'end point' PCR (Figure 3.4), the presence and abundance of a suite of aerobic and facultatively anaerobic genes in the upper 0 – 22 cm region (*e.g.*, *amoA*, *nirS*, and *nirK*) was an initial indication that this region was oxidising and aerobic, despite the high TOC and rapid transition to anaerobic conditions normally expected in such conditions `reference`. Furthermore, the increase in bacterial *amoA* between 0 – 8 cm resulted in the conclusion that 8 cm represented the peak of oxidising conditions in the core. Subsequent 16S rRNA gene taxonomic and Tax4Fun-derived functional profiles corroborated the 'end point' PCR findings, with the ammonia oxidising archaea *Nitrosopumilaceae* and bacteria *Nitrosomonadaceae* increasing in relative abundance between 0 – 8 cm (Figure 3.15). It is not possible to discern the precise mechanism by which oxidising conditions persevered, and in-fact increased between ~ 4 – 12 cm, though groundwater flow in the site is the likely cause (MacKenzie *et al*., 1991), as well as thermally stable, refractory carbon preventing heterotrophic processes occurring.

It was anticipated prior to the project and in initial studies by the BGS that the U in this upper 0 – 22 cm region of the core was likely to be in the insoluble, immobile 4+ state (MacKenzie *et al*., 1991; Jamet *et al*., 1993). However, based on the microbial profiles generated in this project, it was anticipated that any U in the 0 – 22 cm region of the core would instead be in the oxidised U^{6+} species, as more thermodynamically favourable TEAs were likely present in abundance (*e.g.*, ammonia oxidation, denitrification), indicating a lack of direct bioreduction of U in this region. Other metals such as Fe^{3+} and Mn^{4+} are also unlikely to be reduced in this zone for the same reason as U. Indirect reduction of U^{6+} is thought to be unlikely due to the predicted lack of reduced Fe or sulfide species expected in this zone. This prediction of the dominance of U^{6+} species in the upper region was later corroborated *via* XANES which showed that at 10 cm, between 75 – 92% of the bulk U was in the 6+ state (Figure 3.27). The accumulation of U in this region (Figure 3.26) in the normally soluble and mobile U^{6+}, rather than the insoluble and immobile U^{4+}

expected, is counter-intuitive and potentially pertains to the fact U^{6+} can bind to refractory organic ligands (Belli *et al*., 2015).

In the lower half of the core, the emergence of strictly anaerobic and reducing genes, namely *dsrA* and *mcrA* in the 'end point' PCR profiles (Figure 3.4) in tandem with the decrease in the aforementioned genes lead to the initial prediction that any U in this region would be in the reduced U^{4+} state, as it will have been preferentially reduced before sulfate and methanogenesis (Thauer *et al*., 1977). 16S rRNA gene taxonomic and inferred functional profiles provided further evidence of the reducing nature of this region, and the likely reduction of U, with the emergence of *Deltaproteobacteria*, and a suite of methanogenic archaea such as *Methanosaeta* spp., *Methanolinea* spp., and *Methanoregula* spp. (Figure 3.15). Fe^{3+} and Mn^{4+} were also expected to have been reduced in this region, due also to the thermodynamic favourability of these metals as TEAs. The expected presence of reduced Fe^{2+} may also act as a reducing agent for U^{6+} (Du *et al*., 2011). HERFD-XANES later confirmed the predictions based on the microbial profiles of this region that the U will predominantly be reduced, showing between 70 – 90% of the bulk U being the reduced U^{4+} at 30 cm (Table 3.3).

Of particular interest was the 35 cm zone. Early 'end point' PCR profiles showed an increase in bacterial *amoA*, *nirS* and *nirK* and a notable decline in *dsrA* centred around 35 cm (Figure 3.4), despite this region being expected to be reducing and anaerobic. This was initially attributed to the low resolution of the method, but subsequent corroboration *via* 16S rRNA gene sequencing profiles demonstrated that this was probably a real signal in the microbial community. Furthermore, 16S rRNA gene sequencing analysis elucidated the presence of anaerobic methane oxidising *Methylomirabilis* spp., the *mxaF* gene, and *Rokubacteriales* NC10 (Figures 3.14, 3.15, and 3.21D). The disproportionation of O_2 in this process (Ettwig *et al*., 2010) may lead to the introduction of more oxidising conditions in this zone. Therefore, the microbial profiles lead to the prediction of U reoxidation in this zone, due to how readily reoxidised poorly crystalline UO_2 is in the presence of O_2 (Singh *et al*., 2014). HERFD-XANES did indeed show an increase in U^{6+} as a proportion of total U, though the two methods used (*i.e.*, the L_3- and M_4-edge had the greatest discrepancy out of all the samples measured with this method – 27 and 75% as U^{6+} respectively (Table 3.3).

Overall, the range of microbial methods used here were able to accurately anticipate the dominant U redox species at various regions throughout the core, predicting the likelihood of both direct and indirect bioreduction and reoxidation of U within the system. Even the relatively rudimentary, low-resolution 'end point' PCR profiles were able to predict the prevailing geochemical conditions and thus expected U species throughout the core. This highlights the efficacy of using microbial DNA as a biomarker to construct taxonomic and functional profiles that can be used to make direct inferences of the medium-term geochemical conditions of a system.

3.5 References

Alberquerque L, França L, Rainey FA, Schumann P, Fernanda Nobre M, da Costa MS (2011) *Gaiella occulta* gen. nov., sp. nov., a novel representative of a deep branching phylogenetic lineage within the class *Actinobacteria* and proposal of *Gaiellaceae* fam. nov. and *Gaiellales* ord. nov.. Systematic and Applied Microbiology, 34(8): 595–599

Asshauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomics 16S rRNA data. Bioinformatics, 31(7): 2882–2884

Barns SM, Cain EC, Sommerville L, Kuske CR (2007) *Acidobacteria* Phylum Sequences in Uranium-Contaminated Subsurface Sediments Greatly Expand the Known Diversity within the Phylum. Applied and Environmental Microbiology, 73(9): 3113–3116

Blake LI, Halim FA, Gray C, Mair R, Manning DAC, Sallis P, Hutchinson H, Gray ND (2017) Evaluating an anaerobic digestion (AD) feedstock derived from a novel non-source segregated municipal solid waste (MSW) product. Waste Management, 59: 149–159

van Bodegom P, Stams F, Mollema L, Boeke S, Leffelaar P (2001) Methane Oxidation and the Competition for Oxygen in the Rice Rhizosphere. Applied and Environmental Microbiology, 67(8): 3586–3597

Bogard MJ, del Giorgio PA, Boutet L, Garcia Chaves MC, Prairie YT, Merante A, Derry AM (2014) Oxic water column methanogenesis as a major component of aquatic CH_4 fluxes. Nature Communications, 5: 5350

Boyd JA, Woodcroft BJ, Tyson GE (2018) GraftM: a tool for scalable, phylogenetically informed classification of genes within metagenomes. Nucleic Acids Research, gky174, https://doi.org/10.1093/nar/gky174

Brosius J, Palmer ML, Kennedy PJ, Noller HF (1978) Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. Proceedings of the

National Academy of Sciences of the United States of America, 75 (10): 4801-4805

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Pena A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7(5): 335–336

Choppin GR, Wong PJ (1998) The Chemistry of Actinide Behaviour in Marine Systems. Aquatic Geochemistry, 4: 77–101

Concheri G, Stevanato P, Zaccone C, Shotyk W, D'Orazio V, Miano T, Piffanelli P, Rizzi V, Ferrandi C, Squartini A (2017) Rapid peat accumulation favours the occurrence of both fen and bog microbial communities within a Mediterranean, free-floating peat island. Scientific Reports, 7: doi:10.1038/s41598-017-08662-y

Drenovsky RE, Vo D, Graham KJ, Scow KM (2004) Soil Water Content and Organic Carbon Availability Are Major Determinants of Soil Microbial Community Composition. Microbial Ecology, 48: 424–430

Du X, Boonchayaanant B, Wu W-M, Fendorf S, Bargar J, Criddle CS (2011) Reduction of Uranium(VI) by Soluble Iron(II) Conforms with Thermodynamic Predictions. Environmental Science and Tecnology, 45: 4718–4725

Elias DA, Krumholz LR, Wong D, Long PE, Suflita JM (2003) Characterization of Microbial Activities and U Reduction in a Shallow Aquifer Contaminated by Uranium Mill Tailings. Microbial Ecology, 46: 83–91

Emerson D (2012) Biogeochemistry and microbiology of microaerobic Fe(II) oxidation. Biochemical Society Transactions, 40(6): 1211–1216

Ettwig KF, Butler MK, Le Paslier D, Pelletier E, Mangenot S, Kuypers MMM, Schreiber F, Dutilh BE, Zedelius J, de Beer D, Gloerich J, Wessels HJCT, van Alen T, Leusken F, Wu ML, van de Pas-Schoonen KT, Op den Camp HJM, Janssen-Megens EM, Francoijs K-J, Stunnenberg H, Weissenbach J, Jetten MSM, Strous M (2010) Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature, 464: 543–548

Evans PN, Parks DH, Chadwick GL, Robbins SJ, Orphan VJ, Golding SD, Tyson GW (2015) Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. Science, 350(6259): 434–438

Finneran KT, Anderson RT, Nevin KP, Lovley DR (2002) Potential for Bioremediation of Uranium-Contaminated Aquifers with Microbial U(VI) Reduction, Soil and Sediment Contamination: An International Journal, 11(3): 339-357

Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. Proceedings of the National Academy of Sciences of the United States of America, 102(41): 14683-14688

Frank YA, Kadnikov VV, Lukina AP, Banks D, Beletsky AV, Mardanov AV, Sen'kina EI, Avakyan MR, Karnachuk OV, Ravin NV. (2016) Characterization and Genome Analysis of the First Facultatively Alkaliphilic *Thermodesulfovibrio* Isolated from the Deep Terrestrial Subsurface. Frontiers in Microbiology, https://doi.org/10.3389/fmicb.2016.02000

Gorby YA, Lovley DR (1992) Enzymic uranium precipitation. Environmental Science and Technology 1992 ,26(1): 205-207

Gray ND, Head IM (2005) Minerals, mats, pearls and veils: themes and variations in giant sulfur bacteria, in: Gadd GM, Semple KT, Lappin-Scott H (Ed.), Micro-organisms and Earth Systems: 35–71

Guan J, Xia LP, Wang LY, Liu JF, Gu JD, Mu BZ (2013) Diversity and distribution of sulfate-reducing bacteria in four petroleum reservoirs detected by using 16S rRNA and *dsrAB* genes. International Biodeterioration and Biodegradation, 76: 58–66

Hallbeck L, Pedersen K (2014) The Family *Gallionellaceae*, in: Rosenberg E, DeLong EF, Lory S, Strackebrandt E, Thompson F (Ed.), The Prokaryotes (Fourth Edition): 853–858

Harbison AB, Price LE, Flythe MD, Bräuer SL (2017) *Micropepsis pineolensis* gen. nov., sp. nov., a mildly acidophilic alphaproteobacterium isolated from a poor fen, and proposal of *Micropepsaceae* fam. nov. within *Micropepsales* ord. nov. International Journal of Systematic and Evolutionary Microbiology 67: 839-844, doi: 10.1099/ijsem.0.001681

Haroon MF, Hu S, Shi Y, Imelfort M, Keller J, Hugenholtz P, Yuan Z, Tyson GW (2013) Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. Nature, 500: 567–570

Herring JS (2013) Uranium and Thorium Resources. Nuclear Energy: Selected Entries from the Encyclopedia of Sustainability Science and Technology. N. Tsoulfanidis. New York, NY, Springer New York: 463-490

Hooker PJ (1991) The geology, hydrogeology and geochemistry of the Needle's Eye natural analogue site (EUR--13434). Commission of the European Communities (CEC)

Hug LA, Thomas BC, Sharon I, Brown CT, Sharma R, Hettich RL, Wilkins MJ, Williams KH, Singh A, Banfield JF (2016) Critical biogeochemical functions in the subsurface are associatedwith bacteria from new phyla and little studied lineages. Envrionmental Microbiology, 18(1): 159–173

Imachi H, Sakai S, Sekiguchi Y, Hanada S, Kamagata Y, Ohashi A, Harada H (2008) *Methanolinea tarda* gen. nov., sp. nov., a methane-producing archaeon isolated from a methanogenic digester sludge. International Journal of Systematic and Evolutionary Microbiology 58: 294-301, doi: 10.1099/ijs.0.65394-0

Jamet P, Hooker PJ, Schmitt JM, Ledoux E, Escalier des Orres P (1993) Hydrogeochemical modelling of an active system of uranium fixation by organic soils and sediments (Needle's Eye, Scotland). Mineralium Desposita, 28: 66–76

Kämpfer P, Young C-C, Arum AB, Shen F-T, Jäckel U, Rosselló RR, Lai W-A, Rekha PD (2006) *Pseudolabrys taiwanensis* gen. nov., sp. nov., an alphaproteobacterium isolated from soil. International Journal of Systematic and Evolutionary Microbiology 56: 2469-2472, doi: 10.1099/ijs.0.64124-0

Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA, Kuramae EE (2106) The Ecology of *Acidobacteria*: Moving beyond Genes and Genomes. Frontiers in Microbiology, 7: doi:10.3389/fmicb.2016.00744

Kishimoto N, Kosako Y, Tano T (1991) *Acidobacterium capsulatum* gen. nov., sp. nov.: An acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. Current Microbiology, 22(1): 1–7

Klein M, Freidrich M, Roger AJ, Hugenholtz P, Fishbain S, Abicht H, Blackall LL, Stahl DA, Wagner M (2001) Multiple Lateral Transfers of Dissimilatory Sulfite Reductase Genes between Major Lineages of Sulfate-Reducing Prokaryotes. Journal of Bacteriology, 183(20): 6028–6035

Kojima H, Shinohara A, Fukui M (2015) Sulfurifustis variabilis gen. nov., sp. nov., a sulfur oxidizer isolated from a lake, and proposal of Acidiferrobacteraceae fam. nov. and Acidiferrobacterales ord. nov.. International Journal of Systematic and Evolutionary Microbiology, 65: 3709–3713

Konhauser KO, Mortimer RJB, Morris K, Dunn V (2002) The tole of microorganisms during sediment diagenesis: implications for radionuclide mobility, in: Keith-Roach MJ, Livens FR (Ed), Interactions of Microorganisms with Radionuclides: 61–100

Kopylova Y, Guseva N, Shestakova A, Khavschevskaya A, Arakchaa K (2015) Uraniun and thorium behaviour in groundwater of the natural spa area "Choygan mineral water" (East Tuva). IOP Conference Series: Earth and Environmental Science, 27: doi:10.1088/1755-1315/27/1/012034

Lopez-Capel E, Abbott GD, Thomas KM, Manning DAC (2006) Coupling of thermal analysis with quadrupole mass spectrometry and isotope ratio mass spectrometry for simultaneous determination of evolved gases and their carbon isotopic composition. Journal of Analytical and Applied Pyrolysis, 75(2): 82–89

Lovley DR, Phillips EJP, Gorby YA, Landa ER (1991) Microbial reduction of uranium. Letters to Nature, 350: 413–416

Lovley DR, Roden EE, Phillips EJP, Woodward JC (1993) Enzymatic iron and uranium reduction by sulfate-reducing bacteria. Marine Geology, 113: 41–53

Luesken FA, Wu ML, Op den Camp HJM, Keltjens JT, Stunnenberg H, Francoijs K-J, Strous M, Jetten MSM (2012) Effect of oxygen on the anaerobic methanotroph 'Candidatus Methylomirabilis oxyfera': kinetic and transcriptional analysis.

Environmental Microbiology, 14: 1024–1034. doi:10.1111/j.1462-2920.2011.02682.x

Luton PE, Wayne JM, Sharp RJ, Riley PW (2002) The *mcrA* gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. Microbiology, 148: 3521–3530

Ma K, Liu X, Dong X (2006) *Methanosaeta harundinacea* sp. nov., a novel acetate-scavenging methanogen isolated from a UASB reactor. International Journal of Systematic and Evolutionary Microbiology 56: 127-131, doi: 10.1099/ijs.0.63887-0

MacKenzie AB, Scott RD, Houston CM, Hooker PJ (1991) Nautral decay series radionuclide studies at the Needle's Eye natural analogue site (EUR—13126). Commission of the European Communities (CEC)

Markich SJ (2002) Uranium Speciation and Bioavailability in Aquatic Systems: An Overview. The Scientific World, 2: 707–729

McCann CM (2012) The use of natural Mn oxide-containing wastes as a contaminated land remediation strategy and their effects on soil microbial functioning. PhD thesis, Newcastle University, viewed 1 June

McTavish H, Fuchs JA, Hooper AB (1993) Sequence of the gene coding for ammonia monooxygenase in Nitrosomonas europaea. Journal of Bacteriology, 175(8): 2436-2444

Morford JL, Emerson S (1999) The geochemistry of redox sensitive trace metals in sediments. Geochimica et Cosmochimica Acta 63(11-12): 1735-1750

Morris R, Schauer-Gimenez A, Bhattad U, Kearney C, Struble CA, Zitomer D, Maki JS (2014) Methy coenzyme M reductase (*mcrA*) gene abundance correlates with activity measurements of methanogenic H_2/CO_2-enriched anaerobic biomass. Microbial Biotechnology, 7(1): https://doi.org/10.1111/1751-7915.12094

Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology, 59(3): 695–700

Navarrete AA, Kuramae EE, de Hollander M, Pijl AG, van Veen JA, Tsai SM (2012) Acidobacterial community responses to agricultural management of soybean in Amazon forest soils. FEMS Microbial Ecology, 83: 607–621

Nelson WC, Stegen JC (2015) The reduced genomes of Parcubacteria (OD1) contain signatures of a symbiotic lifestyle. Frontiers in Microbiology, 6: doi: 10.3389/fmicb.2015.00713

Nielsen KM, Johnsen PJ, Bensasson D, Daffonchio D (2007) Release and persistence of extracellular DNA in the environment. Environmental Biosafety Research, 6: 37–53

Oude Elferink SJ, Akkermans-van Vliet WM, Bogte JJ, Stams AJ (1999) Desulfobacca acetoxidans gen. nov., sp. Nov., a novel acetate-degrading sulfate reducer isolated from sulfidogenic granular sludge. International Journal of Systematic Bacteriology, 49(2): 345–350

Pester M, Schleper C, Wagner M (2011) The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. Current Opinion in Microbiology, 14(3): 300–306

Prosser JI, Head IM, Stein LY (2014) The family Nitrosomonadaceae, in: Rosenberg E, DeLong EF, Lory S, Strackebrandt E, Thompson F (Ed.), The Prokaryotes (Fourth Edition): 901–918

Prosser JI, Nicol GW (2016) *Candiatus* Nitrosotaleaceae, in: Whitman WB, Rainey F, Kämpfer P, Trujillo M, Chun J, DeVos P, Hedlund B, Dedysh S (Ed.), Bergey's Manual of Systematics of Archaea and Bacteria doi:10.1002/9781118960608.fbm00264

Qin W, Heal K, Ramdasi R, Kobelt J, Martens-Habbena W, Bertagnolli A, Amin S, Walker C, Urakawa H, Könneke M, Devol A, Moffett J, Armbrust E, Jensen G, Ingalls A, Stahl D (2017) *Nitrosopumilus maritimus* gen. nov., sp. nov., *Nitrosopumilus cobalaminigenes* sp. nov., *Nitrosopumilus oxyclinae* sp. nov., and *Nitrosopumilus ureiphilus* sp. nov., four marine ammonia-oxidizing archaea of the phylum *Thaumarchaeota.* International Journal of Systematic and Evolutionary Microbiology, 67(12):5067-5079 doi:10.1099/ijsem.0.002416

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Runde W (2002) The Chemical Interactions of Actinides in the Environment. Los Alamos Science, 26: 392–411

Sakai S, Ehara M, Tseng I-C, Yamaguchi T, Bräuer SL, Cadillo-Quiroz H, Zinder SH, Imachi H (2012) *Methanolinea mesophila* sp. nov., a hydrogenotrophic methanogen isolated from rice field soil, and proposal of the archaeal family *Methanoregulaceae* fam. nov. within the order *Methanomicrobiales*. International Journal of Systematic and Evolutionary Microbiology 62: 1389-1395, doi: 10.1099/ijs.0.035048-0

Schäfer C, Friedrich B, Lenz O (2013) Novel, Oxygen-Insensitive Group 5 [NiFe]-Hydrogenase in *Ralstonia eutropha*. Applied and Environmental Microbiology, 79(17): 5137–5145

Sekiguchi Y, Muramatsu M, Imachi H, Narihiro T, Ohashi A, Harada H, Hanada S, Kamagata Y (2008) *Thermodesulfovibrio aggregans* sp. nov. and *Thermodesulfovibrio thiophilus* sp. nov., anaerobic, thermophilic, sulfate-reducing bacteria isolated from thermophilic methanogenic sludge, and emended description of the genus *Thermodesulfovibrio*. International Journal of Systematic and Evolutionary Microbiology 58: 2541-2548, doi: 10.1099/ijs.0.2008/000893-0

Singh G, Sevinç Şengör S, Bhalla A, Kumar S, De J, Stewart B, Spycher N, Ginn TM, Peyton BM, Sani RK (2014) Reoxidation of Biogenic Reduced Uranium: A Challenge Toward Bioremediation. Critical Reviews in Environmental Science and Technology, 44: 391–415

Søndergaard D, Pedersen CNS, Greening C (2016) HydDB: A web tool for hydrogenase classification and analysis. Scientific Reports, 6: doi:10.1038/srep34212

Stephen JR, Chang YJ, Macnaughton SJ, Kowalchuk GA, Leung KT, Flemming CA, White DC (1999) Effect of toxic metals on indigenous soil β-subgroup proteobacterium ammonia oxidizer community structure and protection against toxicity by inoculated metal-resistant bacteria. Applied and Environmental Microbiology, 65(1): 95-101

Thauer RK, Jungermann K, Decker K (1977) Energy Conservation in Chemotrophic Anaerobic Bacteria. Bacteriological Reviews, 41: 100–180. doi:10.1073/pnas.0803850105

Thauer RK (2012) The Wolfe cycle comes full circle. PNAS, 109(38): 15084–15085

Throbäck IN, Enwall K, Jarvis A, Hallin S (2004) Reassessing PCR primers targeting nirS, nirK and nosZ genes for community surveys of denitrifying bacteria with DGGE. FEMS Microbiology Ecology, 49(3): 401-417

Tribovillard N, Algeo TJ, Lyons T, Riboulleau A (2006) Trace metals as paleoredox and paleoproductivity proxies: An update. Chemical Geology, 232: 12 –32

Vaksmaa A, Guerro-Cruz S, van Alen TA, Cremers G, Ettwig KF, Lüke C, Jetten MSM (2017) Enrichment of anaerobic nitrate-dependent methanotrophic '*Candidatus* Methanoperedens nitroreducens' archaea from an Italian paddy field soil. Applied Microbiology and Biotechnology, 101(18): 7075–7084

Vignais PM, Billoud B (2007) Occurrence, Classification, and Biological Function of Hydrogenases: An Overview. Chemical Reviews, 107(10): 4206–4272

Weiss JV, Rentz JA, Plaia T, Neubauer SC, Merrill-Floyd M, Lilburn T, Bradburne C, Megonigal JP, Emerson D (2007) Characterization of Neutrophilic Fe(II)-Oxidizing Bacteria Isolated from the Rhizosphere of Wetland Plants and Description of *Ferritrophicum radicicola* gen. nov. sp. nov., and *Sideroxydans paludicola* sp. nov.. Geomicrobiology Journal, 7(8): 559–570

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y (2006) *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum *Chloroflexi*. International Journal of Systematic and Evolutionary Microbiology 56: 1331-1340, doi: 10.1099/ijs.0.64169-0

Chapter 4. Using Microbial Community and Function to Understand and Predict Biogeochemical Cycling and Radionuclide Fate in Two Sellafield-Impacted Sediment Cores from the Eastern Irish Sea

4.1 Introduction

Since Sellafield (Cumbria, UK) was first converted into a civil nuclear power station in the 1950s, atmospheric, and later aqueous, discharges of low-level radionuclides into the adjacent Irish Sea has occurred, including the actinides U, Np, Pu, and Am and other radionuclide fission products such as Cs and Sr (Kershaw *et al.*, 1992; Gray *et al.*, 1995). These radionuclide contaminants accumulate in local coastal areas where they are deposited at depth by sedimentation (Kershaw *et al.*, 1990; MacKenzie and Scott, 1993). This sedimentation has created a profile of radionuclide discharge levels from Sellafield within the sediments (Ray *et al.,* in prep.). These discharges and sedimentation profiles of radionuclides from Sellafield and the Irish Sea and coast create a unique natural laboratory to investigate the long-term interaction between radionuclides, the sediments in which they are deposited, and the indigenous microbial communities which mediate geochemical processes in these marine environments. The Irish Sea is a macro-tidal, semi-enclosed marine system (Halcrow Group, 2013) connected to the Atlantic Ocean from both the north *via* the North Channel and the south *via* St. George's Channel (Carr and Blackley, 1986).

A proportion of the aqueous discharges from Sellafield since the 1950s have been deposited and have become associated with a transect of finely-grained sediments on the Irish Sea bed, referred to as the Irish Sea Mud Patch (Hetherington and Harvey, 1978). Aqueous discharges of radionuclides from Sellafield are carried out by the sea and tide to the outer extent of the tide at the site of the Mud Patch. Here, radionuclides are not deposited at depth as would be expected with normal sedimentation but are instead mixed by *polychaete* worms such as *Maxmulleria lankesteri* which burrows up to 40 cm into the Mud Patch sediment, reworking and redistributing the radionuclides into the water column (Kershaw *et al.*, 1983). These radionuclides are then transported elsewhere by tidal currents, contaminating the coastlines of the UK and Ireland (McCartney *et al.*, 1994; Keogh *et al.*, 2007; Al-Qasmi *et al.*, 2016 and 2017). This is the foundation of the source-sink relationship between the Irish Sea Mud Patch and the River Esk estuary at Ravenglass. The range of radionuclides released from Sellafield into the Irish Sea and found in the

Mud Patch sediments include isotopes of caesium (Cs), plutonium (Pu), americium (Am), among many others (McCartney *et al.*, 1994).

By understanding the redox conditions in radionuclide impacted sediments, there is a potential benefit in creating taxonomic and metabolic microbial functional profiles from the depth differentiated DNA to enable inferences about the prevailing geochemical and redox conditions of an environment to be made. This is because identification of the presence, abundance and differential spatial distribution of key monophyletic taxonomic clades that have specific nutrient and environmental requirements must reflect the presence of the stable geochemical conditions required to enable growth of those taxa. A natural environment with sufficient labile carbon and oxygen levels would be expected to be dominated by heterotrophic aerobic microbial species, as this is the most favourable respiratory process available. likewise, the presence of microbial clades with specific environmental requirements, such as chemolithoautotrophic ammonia oxidising, or chemoheterotrophic sulfate-reducing bacteria, imply that conditions of the surrounding environment were such that favoured the growth of these microbial species. Furthermore, where a system has a significant population of one of these monophyletic clades clearly restricted to a particular depth and bounded above and below by taxa assignable to a different respiratory function, this patterning implies that this respiratory function must have predominated at this depth for a substantial period of time, *i.e.*, weeks to months if not years consistent with the timeframes for microbial growth and persistence. Another benefit of this taxonomic and functional approach is not just the broad identification of redox conditions but also the identification within these zones of tightly coupled oxidative and reductive biogeochemical cycles co-occurring spatially.

This study used a range of microbial molecular techniques to create profiles of microbial taxonomic and functional abundances as a function of depth from two cores contaminated with Sellafield-derived radionuclides. These profiles were used to detect depth resolved spatial distribution of microbial communities and function and identify any depth-related distribution changes of specific key microbial taxonomic and functional assemblages that may indicate the prevailing geochemical conditions within the core. Microbial community assemblages were used to make inferences about the long-term transport and fate of redox-sensitive radionuclides.

4.2 Methods

Full methods are described in the methods and materials chapter (Chapter 2) of this thesis. A brief description of the methods used in this particular chapter are provided here.

4.2.1 Sample collection and preparation

One sediment core was collected from a belt of fine-grained sediment in the north-east Irish Sea mud-patch (hereby referred to as Mud Patch), (54°26'80" N, 03°42'89" W) in June 2014, and one sediment core from the intertidal zone of the Ravenglass saltmarsh (54°20'24" N, 03'24'06" W) in September 2014. See Chapter 2.1.2.

4.2.2 DNA Extraction and Quantification

DNA was extracted from each core separately, but the method and kits used were identical. DNA was extracted from ~ 500 mg of soil for each sample, and the FastDNA Spin Kit for Soil (MP Biomedicals, USA) was used as per manufacturer's instructions. For the purposes of performing statistical analyses on the data generated on samples within each core, samples from the two cores were divided into four arbitrary groups. Ravenglass was divided into four groups: Group 1, 1 – 5 cm; Group 2, 6 – 10 cm; Group 3, 12 – 20 cm; Group 4, 22 – 28 cm. Mud Patch was divided into five groups: Group 1, 1 – 5 cm, Group 2, 6 – 10 cm; Group 3, 12 – 20 cm; Group 4, 22 – 30 cm; Group 5, 32 – 36 cm. See Chapter 2.2.

4.2.3 'End Point' PCR Functional Gene Series

A series of primer pairs were used to target a range of functional, redox-related genes (Table 3.2). Full details of target genes and procedural positives are in the methods chapter (Tables 2.1 and 2.2). See Chapter 2.3.

4.2.4 16S rRNA Gene Sequencing Preparation

All samples including the procedural negative were prepared for Ion Torrent sequencing from the control core. A multiplexed FASTQ file containing all the sequenced reads was returned from sequencing. See Chapter 2.5.1.1.

4.2.5 16S rRNA Gene Sequencing Pipeline Analysis

Pipeline analysis of sequencing data generated by both Ion Torrent and Illumina 16S rRNA gene sequencing was performed in QIIME 2 (version 2018.6) (Caporaso *et al.*, 2010), using a bespoke set of scripts, which can be found at https://github.com/peterleary/q2pipeline.

4.2.6 Tax4Fun Analysis

The closed-reference OTU Table produced by the pipeline was used as inputs into the Tax4Fun package in R (Asshauer *et al.*, 2015) to create a profile of inferred abundances of KEGG Ortholog genes based on OTU taxonomy. A predefined set of genes of interest were targeted in the Tax4Fun profile (Table 2.4). See Chapter 2.6.

4.2.7 Geochemical Analysis

Sequential Loss on Ignition (LOI), ICP-AES, XRF α-spectrometer were used to measure TOC, porewater metals and anions, solid phase metals, and plutonium activity. See Chapter 2.9 and 2.9.2.

4.3 Results

4.3.1 DNA Extractions

The volume of DNA extracted from the Ravenglass depth core samples decreased with increasing depth (Figure 4.1A). Extracted DNA volumes ranged from a peak of 28.6 ng/μL at 4 cm, and a low of 1.22 ng/μL at 28 cm. Between 1 – 9 cm, the average volume was 16.54 ± 2.09 ng/μL, and between 10 – 28 cm, the average volume was 2.14 ± 0.3 ng/μL. One-way analysis of variance (AOV) showed a statistical difference in extracted DNA volumes across all sample groups (p value = < 0.05), and Tukey HSD showed statistical differences in concentration between all

sample group pairs with the exception of group pair 3 and 4 (depth groups 12 – 20 cm and 22 – 28 cm, green and purple points on Figure 4.1A) (Appendix D.1A). The volume of DNA extracted from the mud patch depth core was lower than extracted from Ravenglass and showed no discernible pattern with depth. Volumes of extracted DNA from the Mud Patch ranged from 3.98 ng/μL at 4 cm to 0.91 ng/μL at 32 cm. Average DNA extracted from the whole core was 2.21 ± 0.18 ng/μL (Figure 4.2A). AOV did not show a statistical difference in DNA concentrations across the five Mud Patch sample groups (Appendix D.2A).

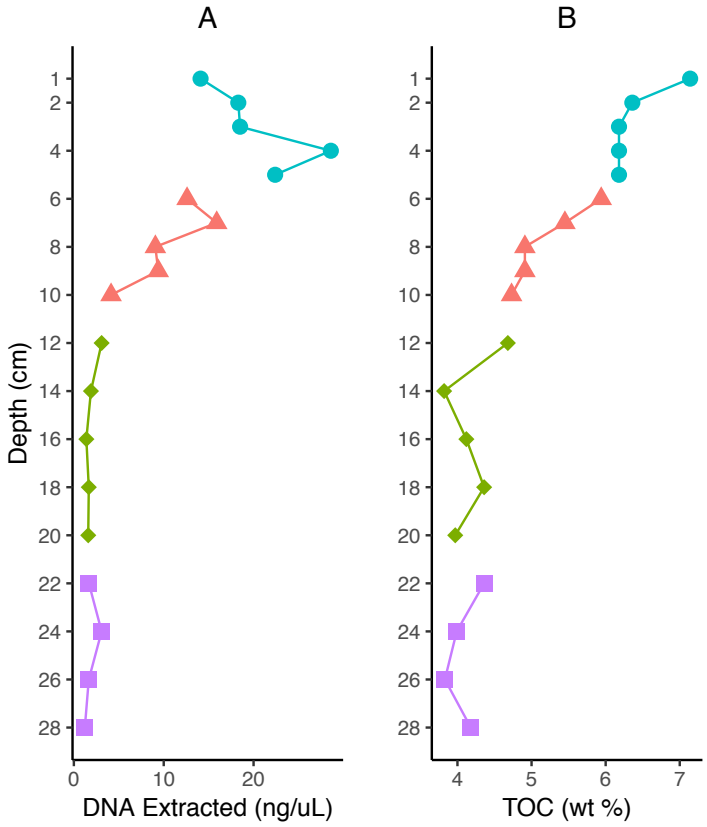


Figure 4.1. Ravenglass Depth Core: (A). The amount of DNA extracted from each depth sample from the core in ng/μL, and; (B). The total organic carbon (TOC) measured by Loss on Ignition (LOI) for the depth samples measured. Point colours and shapes denote sample depth group: blue circles = group 1 (1 – 5 cm); red triangles = group 2 (6 – 10 cm); green diamonds = group 3 (12 – 20 cm), and; purple squares = group 4 (22 – 28 cm)

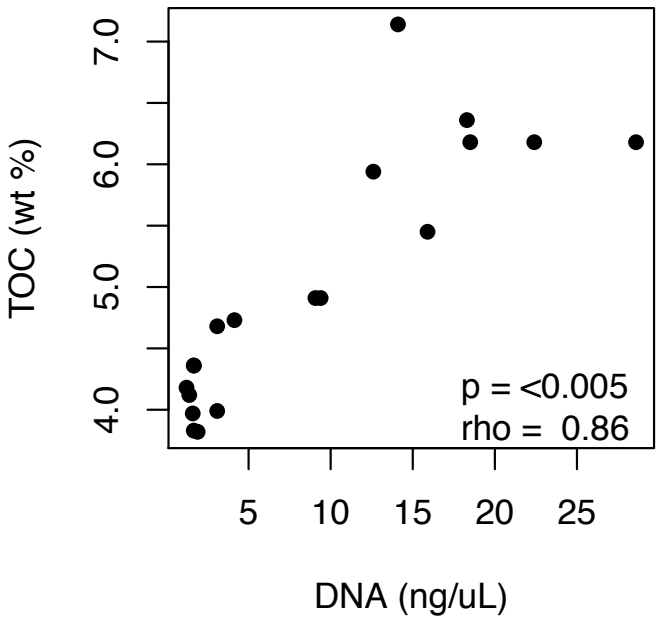


Figure 4.2. Ravenglass Depth Core: XY scatterplot showing the amount of DNA extracted from each sample against TOC value for the same sample. Spearman correlation is presented as ρ and *p* values.

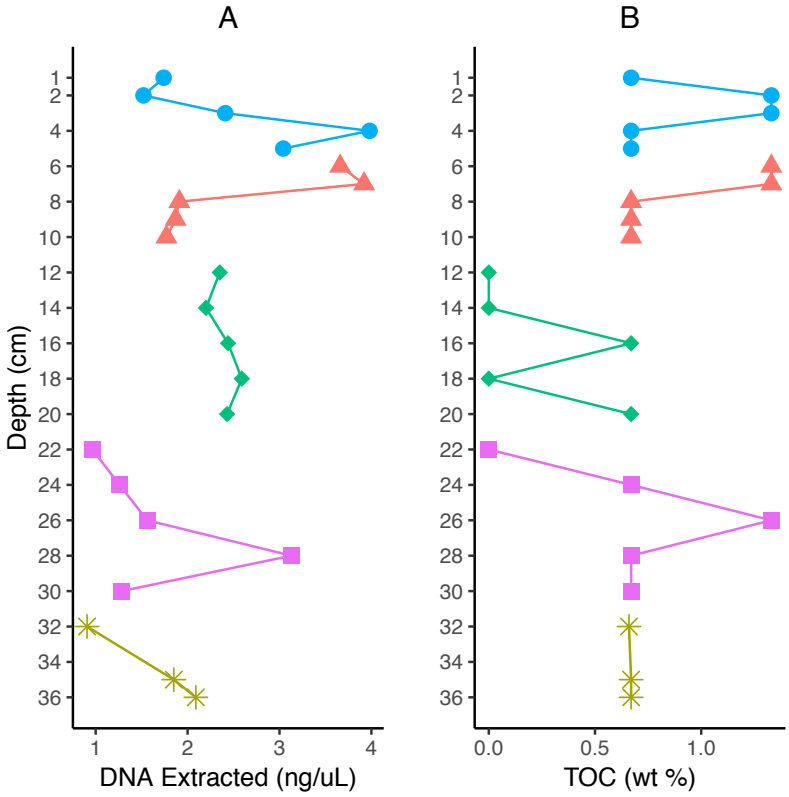


Figure 4.3. Mud Patch Depth Core: (A). The amount of DNA extracted from each depth sample from the core in ng/μL, and; (B). The total organic carbon (TOC) measured by Loss on Ignition (LOI) for the depth samples measured. Point colours and shapes denote sample depth group: blue circles = group 1 (1 – 5 cm); red triangles = group 2 (6 – 10 cm); green diamonds = group 3 (12 – 20 cm); purple squares = group 4 (22 – 30 cm); gold stars = group 5 (32 – 36 cm).

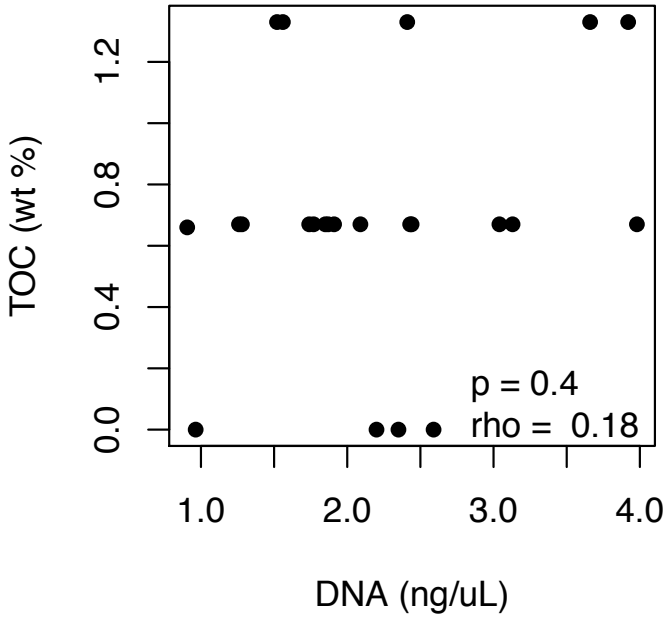


Figure 4.4. Mud Patch Depth Core: XY scatterplot showing the amount of DNA extracted from each sample against TOC value for the same sample. Spearman correlation is presented as ρ and *p* values.

4.3.2 Total Organic Carbon

Total organic carbon (TOC) was measured by Loss on Ignition (LOI) by the University of Manchester. TOC in Ravenglass decreased with depth, ranging from 7.14 (wt %) at 1 cm, to 3.82 (wt %) at 14 cm (Figure 4.1B), and averaged 5.92 ± 0.24 between 1 – 9 cm and 4.2 ± 0.1% between 10 – 28 cm. AOV showed a statistical difference in TOC values across the four sample groups, and Tukey HSD showed a statistical difference between all sample group pairs with the exception of group pair 3 and 4 (Appendix D.1B). Similar to the DNA amounts, TOC in the Mud Patch was lower than in Ravenglass and had no discernible pattern with depth. It did not exceed 1.33%, and averages 0.7 ± 0.09% throughout the core (Figure 4.3B). AOV resulted in a borderline-significant *p* value of 0.063, with Tukey HSD showing borderline-significant *p* values between sample group pair 1 – 3 (*p* value = 0.07), and pair 2 – 3 (*p* value = 0.071) (Appendix D.2B).

That the TOC in the Mud Patch is comprised of only three different figures, 0%, 0.67%, and 1.33%, is because the mass loss at 375˚C was only measured as either 0.00, 0.01, or 0.02 g, and with all samples starting with 1.5 g of mass, there were only three values possible. Measuring to more than two decimal places would have resulted in more specific TOC values being calculated.

Scatterplots and Spearman correlation of extracted DNA as a function of TOC were produced for each core. As DNA volume and TOC both decreased with depth in the Ravenglass core, Spearman correlation resulted in a ρ of 0.86 (p-value = < 0.005) (Figure 4.2). However, in the Mud Patch core, there was no correlation with depth either for, or between, the DNA and TOC. This was reflected in the Spearman ρ of 0.18 (p-value = 0.4) (Figure 4.4).

4.3.3 'End Point' PCR Functional Gene Series

'End point' PCR band intensities for functional genes were quantified *via* BioNumerics, and bands given a grey-scale 8-bit value (pixel value) based on the band size, thickness. Band intensities as defined by the pixel value are presented as a qualitative inference of gene abundance within each depth sample, *i.e.*, a higher band intensity denotes a higher gene abundance within a specific sample.

4.3.3.1 *Ravenglass Depth Core*

The archaeal and bacterial ammonia monooxygenase (*amoA*) gene abundances as inferred from band intensities had different abundances with depth. Archaeal ammonia oxidisers (AOA) (Figure 4.5A) had a relatively low abundance, ranging from a band intensity of 41 at 26 cm to 123 at 18 cm, and did not appear to have a pattern with depth. The average band intensity for AOA in the Ravenglass core was 79.37 ± 4.95. AOV did not show a statistical difference in AOA abundance across the four sample groups (Appendix D.3A). The bacterial ammonia oxidisers (Figure 4.5B) (AOB) increased with depth from 1 – 10 cm, from 108 to 189 (average 149.7 ± 10.13) and decreased between 12 – 28 cm from 166 to 88 (average 120.9 ± 9.2). AOV did show a statistical difference across the four sample groups for AOB band intensity, with Tukey HSD showing a statistical difference between sample group pairs 1 – 2, 2 – 4, and 3 – 4 (Appendix D.3B) The denitrifying *nirK* gene was most abundant in the upper half of the core (Figure 4.5C). Between 1 – 10 cm, the average abundance of *nirK* was 182.4 ± 9.48. This dropped to an average of 53.8 ± 7.05 between 12 – 28 cm. AOV showed a statistical difference across all sample groups for *nirK* band intensity, with Tukey HSD showing a statistical difference between all sample group pairs with the exception of pairs 1 – 2 and 3 – 4 (Appendix D.3C). The other denitrifying gene *nirS* did not have the same pattern with depth as *nirK*, having a more complicated pattern with depth (Figure 4.5D). This gene

increased between 1 – 4 cm from 128 to 164 (average 132.5 ± 14.25), decreased to 86 at 5 cm, and increased between 6 – 8 cm to a peak of 254. Between 9 – 28 cm, there was a general decrease from 231 at 9 cm to 78 at 28 cm (average 133.8 ± 15.97). AOV did not show a statistical difference in *nirS* band intensity across sample groups (Appendix D.3D). Lastly, the sulfate-reducing *dsrA* gene appeared to be most abundant in the upper half of the core (Figure 4.5E). It increased between 1 – 4 cm from 196 to 247 (average 207.3 ± 16.26), dropped from 5 – 7 cm from 231 to 116 (average 176.3 ± 33.32), increased at 8 cm to 196, and between 9 – 28 cm, averaged 78.6 ± 8.34). AOV showed a statistical difference in *dsrA* band intensity across all sample groups, and Tukey HSD showed a statistical difference between all sample group pairs, except pair 3 – 4 (Appendix D.3E). There were no PCR amplicon bands detected in the *mcrA* gene reactions.

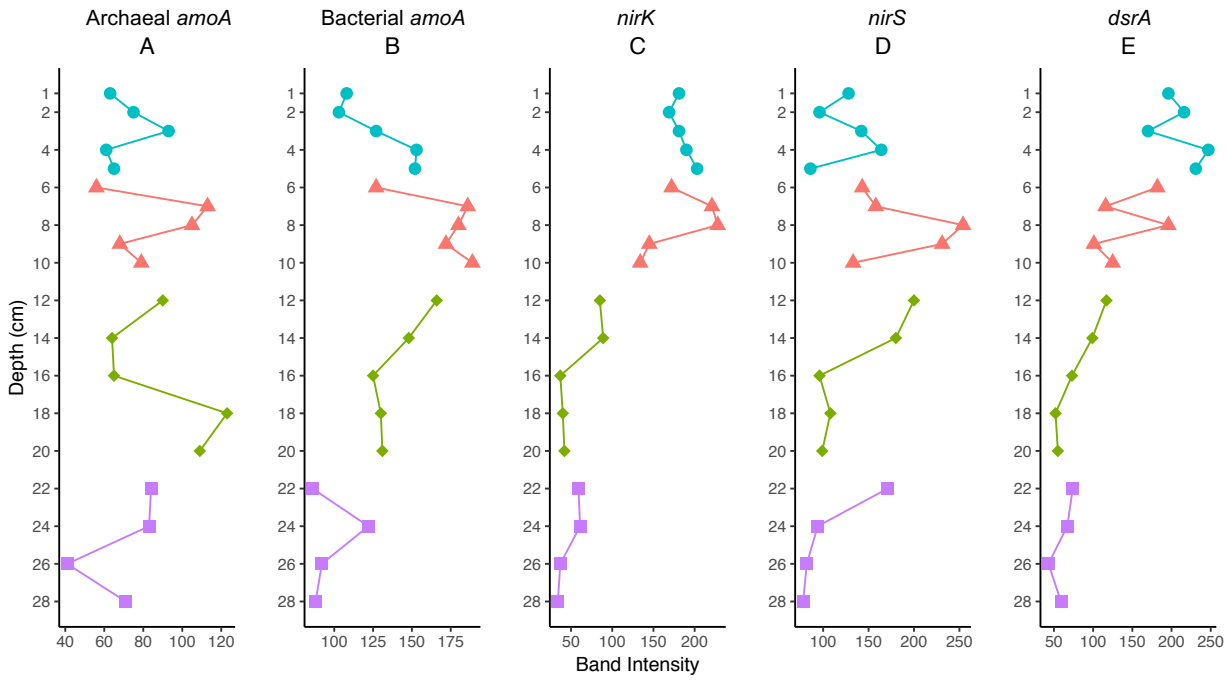


Figure 4.5. Functional gene end point PCR band intensities for the Ravenglass core. Values on the Y axis represent depth in cm, values on the X axis represent the band intensity of the PCR amplicon as calculated by BioNumerics. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA*.

4.3.3.2 *Irish Sea Mud Patch Depth Core*

Similar to the Ravenglass core, the AOA group had a relatively low band intensity throughout the Mud Patch core (Figure 4.6A). AOA abundance increased

from 1 – 5 cm, from 61 to 89 (average 71.2 ± 4.75), between 6 – 12 cm the abundance dipped from 90 at 6 cm to 93 at 14 cm, and a low of 79 between 9 – 10 cm. Between 16 – 28 cm, AOA abundance decreased from 83 to 48 (average 67.7 ± 4.21). From 30 – 36 cm, abundance increased slightly to an average of 56 ± 2.42). AOV showed a statistical difference in AOA band intensity across the five Mud Patch sample groups, with Tukey HSD showing a statistical difference between sample group pairs 2 – 4, 2 – 5, 3 – 4, and 3 – 5 (Appendix D.4A). The AOB appeared to persist throughout the majority of the core (Figure 4.6B), averaging 253.4 ± 0.53 between 1 – 28 cm. At 30 cm, this lowered to 162, before increasing again to 232 at 32 cm (and 250 and 252 at 34 and 36 cm respectively). AOV did not show a statistical difference in AOB band intensity across sample groups (Appendix D.4B).

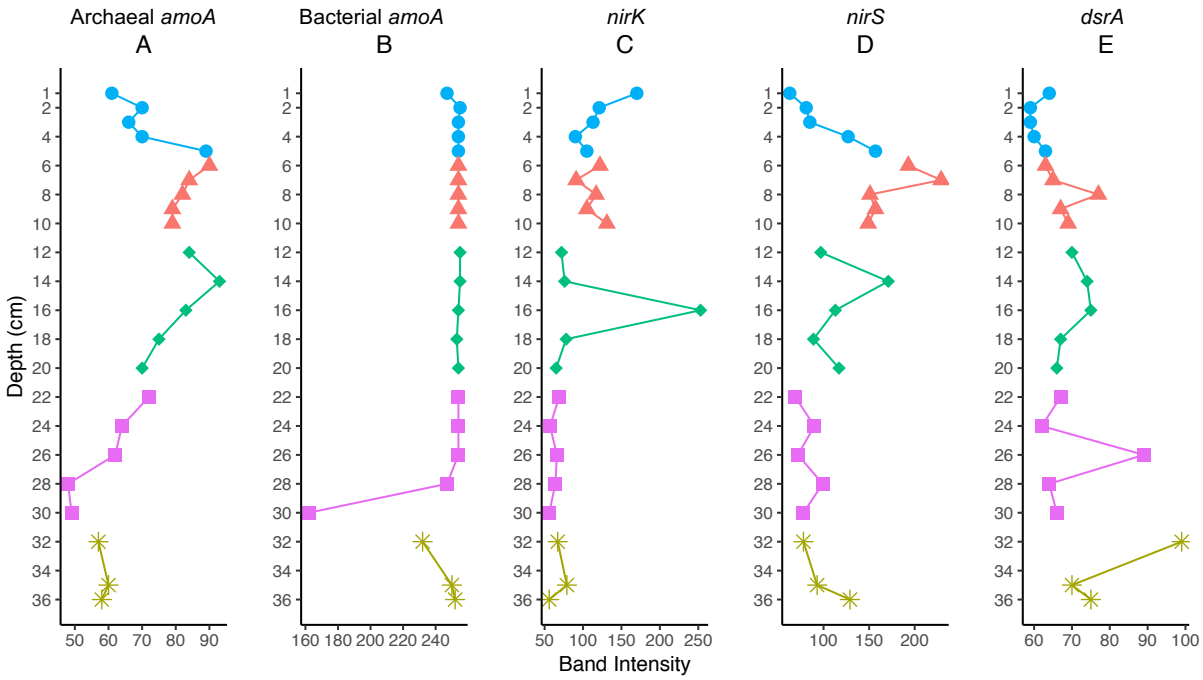


Figure 4.6. Functional gene end point PCR band intensities for the Mud Patch core. Values on the Y axis represent depth in cm, values on the X axis represent the band intensity of the PCR amplicon as calculated by BioNumerics. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA*.

The *nirK* gene abundance (Figure 4.6C) was 170 at 1 cm and decreased to an average of 103.9 ± 5.85. At 16 cm, this increased to 253, and between 18 – 36 cm decreased to an average of 65.7 ± 2.6. AOV did not show a statistical difference in *nirK* band intensity across sample groups (Appendix D.4C). The *nirS* (Figure 4.6D) increased from 1 – 7 cm from 63 to 229 (average 133.6 ± 23.57), then between 8 –

36 cm varied between a peak of 171 at 14 cm and a low of 69 at 22 cm (average between 8 – 36 cm is 109.5 ± 8.2). AOV did show a statistical difference in *nirS* band intensity across sample groups, and Tukey HSD showed a statistical difference between sample group pairs 1 – 2, 2 – 4, and 2 – 5, with group pair 2 – 3 having a *p* value of 0.051 (Appendix D.4D). Lastly, the *dsrA* (Figure 4.6E) was relatively lower in the mud patch than in Ravenglass depth core, averaging 69.1 ± 1.97. AOV showed a statistical difference in *dsrA* band intensity across sample groups, with Tukey HSD showing a statistical difference between sample group pair 1 – 5 (Appendix D.4E). There were no PCR amplicon bands detected in the *mcrA* gene reactions.

4.3.4 16S rRNA Gene Sequencing Results

Sequencing produced 1,089,823 reads in the demultiplexed Ravenglass library, and 812,697 reads in the demultiplexed Mud Patch library. Mud patch depth sample representing 10–12 cm failed to generate any reads during sequencing, and so was excluded from subsequent 16S rRNA gene analysis. A total of 2,432 ASVs from 366,611 reads were identified by DADA2 during pipeline analysis in the Ravenglass core, and 2,380 ASVs from 416,668 reads were identified in the Mud Patch core. ASV tables were rarefied to 13,530 sequences per depth sample for the Ravenglass depth core, and 8,600 for the Mud Patch depth core for core phylogenetic metric diversity analysis in QIIME2. A total of 615 ASVs at genus level from Ravenglass and 598 ASVs at genus level from Mud Patch were identified when taxonomy was collapsed.

4.3.4.1 *Alpha Diversity*

In the Ravenglass depth core, Shannon species richness was similar between sample groups 1 – 3 (1 – 20 cm) and decreased in group 4 (22 – 28 cm) (Figure 4.7A). AOV showed a statistical difference in Shannon index values across all sample groups, and Tukey HSD showed a statistical difference between sample group pair 2 – 4 (Appendix D.5A). Faith phylogenetic diversity appeared to decrease with depth based on the increasing length of the box plot whiskers (Figure 4.7C), with AOV resulting in a borderline-significant *p* value of 0.063, and Tukey HSD showing a statistical difference in Faith indices between sample group pair 2 – 4 (Appendix D.5C). Pielou species evenness increased from sample group 1 to 2, then decreased from 2 through 4 (Figure 4.7E). AOV showed a statistical difference in Pielou index

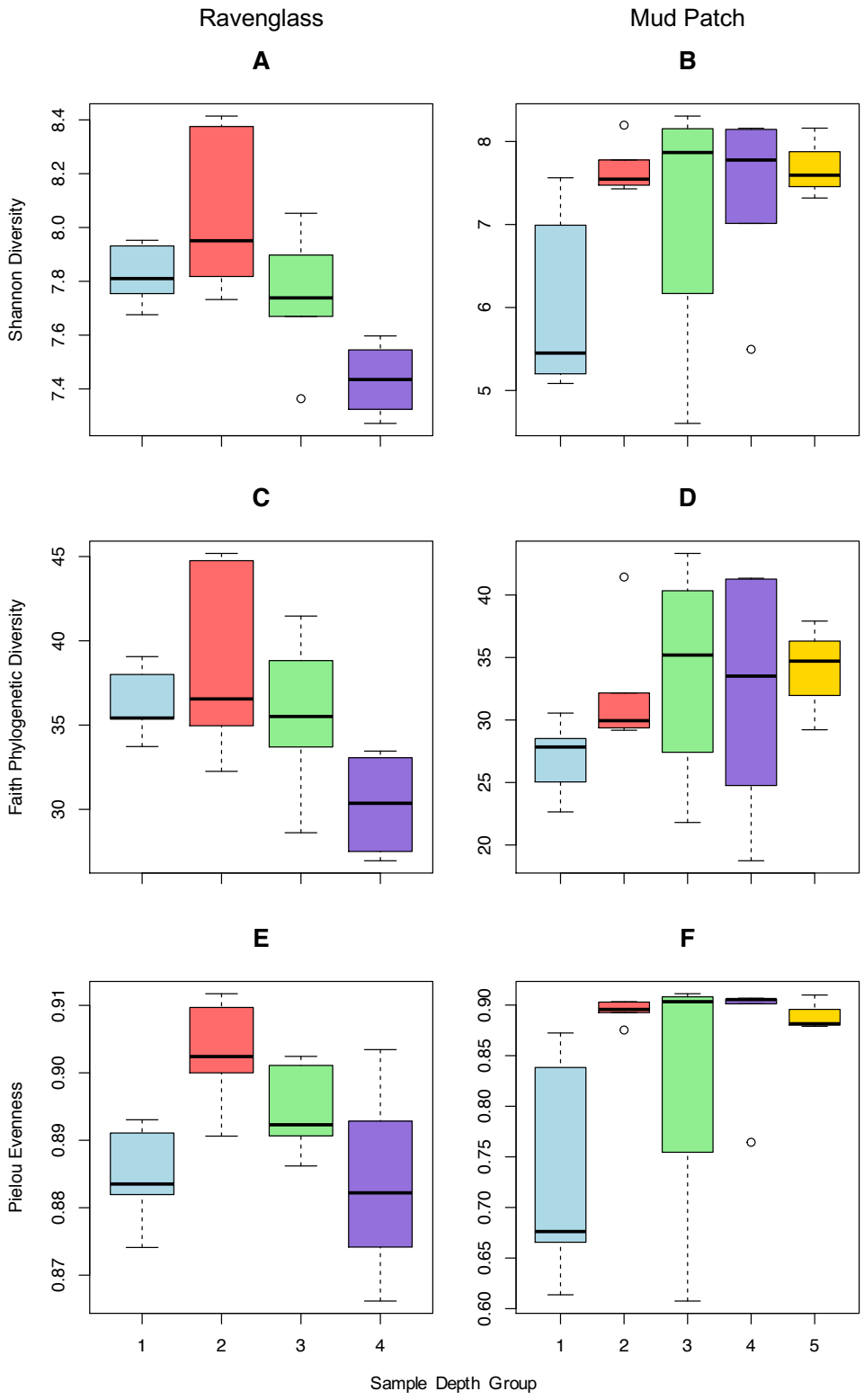


Figure 4.7. Whisker box plots of three alpha diversity indices from the Ravenglass and Mud Patch depth cores. Boxes denote samples divided into arbitrary groups in depth clusters for comparison. Group A = 0 – 5 cm; Group B = 6 – 10 cm; Group C = 12 – 20 cm; Group D = 22 – 28 cm for Ravenglass, and 22 – 30 cm for Mud Patch; Group E = 32 – 36 cm, Mud Patch only. (A). Boxplot of Shannon species richness of Ravenglass; (B). Shannon richness for Mud Patch; (C). Faith phylogenetic diversity for Ravenglass; (D). Faith phylogenetic diversity for Mud Patch; (E). Pielou species evenness for Ravenglass; (F). Pielou species evenness for Mud Patch.

values across all sample groups, and Tukey HSD shows a statistical difference between sample group pair 2 – 4, with pair 1 – 2 having a borderline-significant *p* value of 0.052 (Appendix D.5E). In the Mud Patch depth core, none of indices (Shannon species richness (Figure 4.7B), Faith phylogenetic diversity (Figure 4.7D), nor Pielou species evenness (Figure 4.7F) appeared to have any depth-related patterns, and AOV did not show a statistical difference in Shannon richness of Faith phylogenetic index values across sample groups (Appendix D.5B, 7D), with Pielou evenness AOV resulting in a borderline-significant *p* value of 0.066 (Appendix D.5F).

4.3.4.2 *Beta Diversity*

The microbial community assemblages of both Ravenglass and Mud Patch depth cores were compared together prior to their analysis separately. ASV assemblages of samples from the two cores were compared using all beta diversity ordination methods (PCoA, nMDS, and DCA), however, the ASV community assemblages from the two cores were found to be completely dissimilar across all methods used. Therefore, taxonomic classification of ASVs collapsed at three different levels were compared instead, to compare taxonomic assemblages between the two cores at different levels. Both nMDS and DCA (Figures 4.8A and 4.8B) plots of taxonomic assemblages at species level showed similarities in the taxonomic communities of the two depth cores, with the Ravenglass depth core consistently showing depth-related distribution of samples and the Mud Patch showing a more homogenous community with the exception of several outliers (*e.g.*, 1, 4, 5, 18, and 22 cm).

Looking at the ASV community assemblages of individual cores, PCoA of samples from the Ravenglass depth core only showed a consistent depth-related distribution from 1 cm to 28 cm, across all three dissimilarity methods used (Bray Curtis, and weighted and unweighted UniFrac) (Figure 4.9A, 9B, 9C). However, the large number of zeroes in the ASV table – that is, a large number of ASVs that are present and abundant in some samples and absent in others, *i.e.* sparse dataset – resulted in a ‘arch’ effect on the Ravenglass PCoAs. This ‘arch’ effect is a result of the restriction of many ASVs to a few sample depths, and therefore many zeroes in the ASV table, a sparse table, and results in a loss of community dissimilarity resolution after a certain point of dissimilarity, known as saturation, and that samples at either end of the horseshoe/arch are so dissimilar that it cannot be reliably reflected using the ordination techniques (Morton *et al.*, 2017). PERMANOVA

showed a statistical difference in distances across the four sample groups for all three similarity methods, and pairwise PERMANOVA showed a statistical difference across all sample group pairs with the exception of group pair 3 – 4, across all three methods (Appendix D.6A–C).

Both nMDS (Figure 4.10A) and DCA (Figure 4.10B) plots showed identical patterns of depth-related distribution of samples. The arch effect was clearly negated by the nMDS analysis while still indicating the strict relationship between depth and community composition. The precise reason nMDS does not result in the arch effect is unclear. Hierarchical clustering of samples on the nMDS based on similarity resulted in four main clusters associated with depth.

PCoA of samples from the Mud Patch core did not have any discernible linear depth-related distribution of samples as observed in the Ravenglass depth core, however, clustering of samples from deeper within the sediment versus surface samples was observed (Figure 4.11A, 11B, 11C). All samples from 8 – 36 cm, with the exception of 18 and 22 cm, clustered together based on Bray Curtis and weighted UniFrac similarity (Figure 4.11A and 11B respectively), whereas the was no such defined clustering in unweighted UniFrac (Figure 4.11C). PERMANOVA showed a statistical difference in distance across all five sample groups for Bray Curtis similarity, and pairwise PERMANOVA showed a statistical difference between sample group pairs 1 – 2, 1 – 4, 1 – 5, 2 – 4, and 2 – 5 (Appendix D.7A). Weighted UniFrac did not have a statistical difference in distances (Appendix D.7B). Unweighted UniFrac did have a statistical difference across all five sample groups, and pairwise PERMANOVA showed a statistical difference between all sample group pairs, except pairs 3 – 4, 3 – 5, and 4 – 5 (Appendix D.7C).

Similar to PCoA, nMDS (Figure 4.12A) and DCA (Figure 4.12B) plots showed clustering of samples 8 – 36 cm, excluding 18, 22, and 30 cm, which were outliers. Surface samples 1 – 7 cm formed a second cluster. Hierarchical clustering based on similarity showed four main clusters – 1, 2, 3, 6 cm; 4, 5, 7 cm; 18 and 22 cm, and; 8 – 36 cm, with 30 cm not part of a cluster.

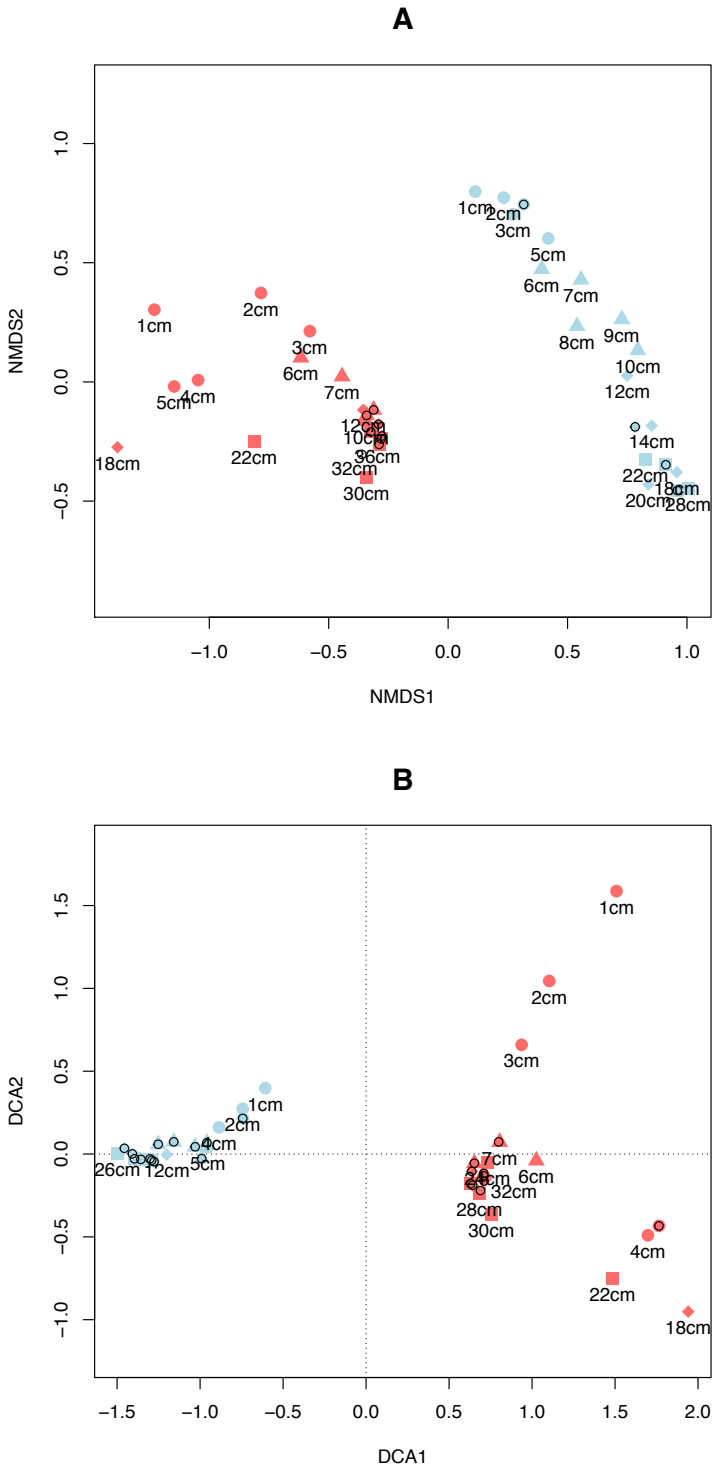


Figure 4.8. Ordination of the ASV taxonomic classification assemblages collapsed at species classification level of both depth core samples. Red samples denote Mud Patch depth core and blue samples denote Ravenglass depth core. Points refer to sample group depth. (A). Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity of taxonomy collapsed at species level; (B). Detrended Correspondence Analysis (DCA) of taxonomy collapsed at species level.

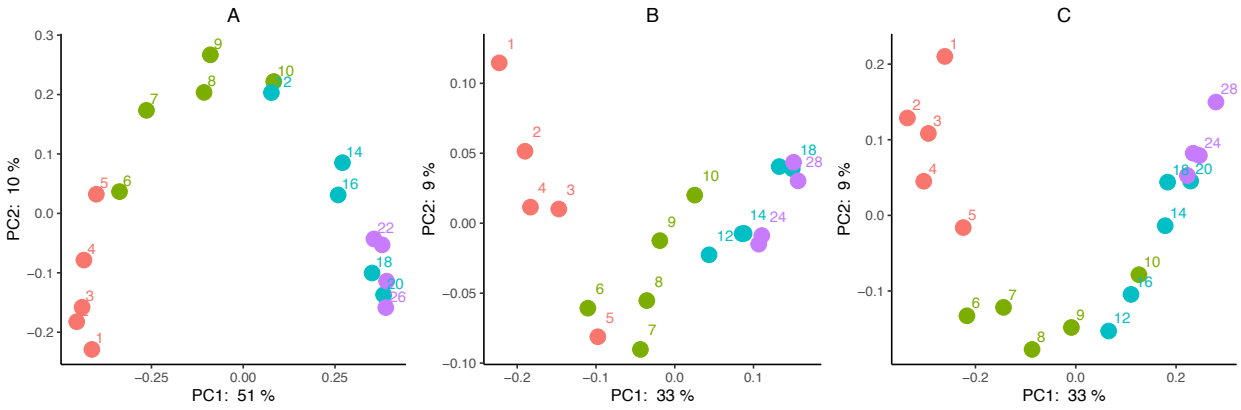


Figure 4.9. Ravenglass Depth Core: PCoA of the microbial community assemblages. Sample colours denote grouping of depth samples. Red = Group 1 (0 – 5 cm); Green = Group 2 (6 – 10 cm); Blue = Group 3 (12 – 20 cm); Purple = Group 4 (22 – 28 cm). Labels represent depth in cm. (A). Bray Curtis similarity; (B). Weighted UniFrac similarity; (C). Unweighted UniFrac similarity.

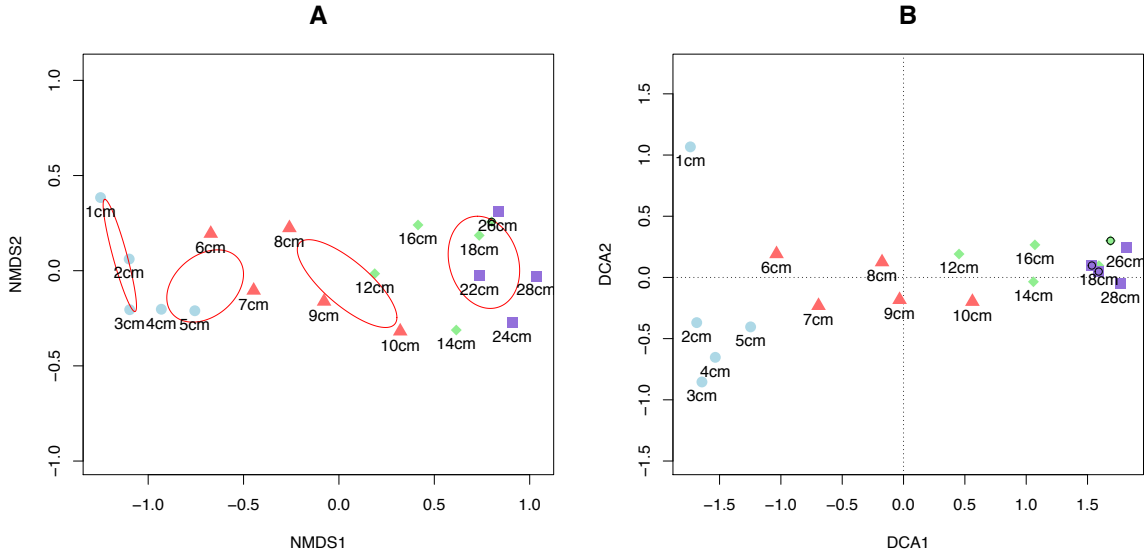


Figure 4.10. Ravenglass Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups.

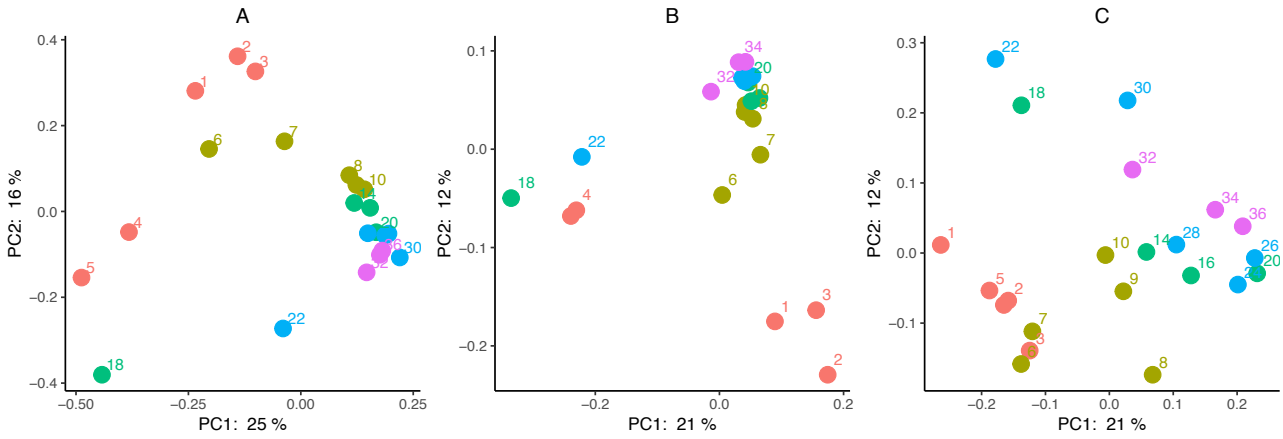


Figure 4.11. Mud Patch Depth Core: PCoA of the microbial community assemblages. Sample colours denote grouping of depth samples. Red = Group 1 (0 – 5 cm); Gold = Group 2 (6 – 10 cm); Green = Group 3 (12 – 20 cm); Blue = Group 4 (22 – 30 cm); Purple = Group 5 (32 – 36 cm). Labels represent depth in cm. (A). Bray Curtis similarity; (B). Weighted UniFrac similarity; (C). Unweighted UniFrac similarity.

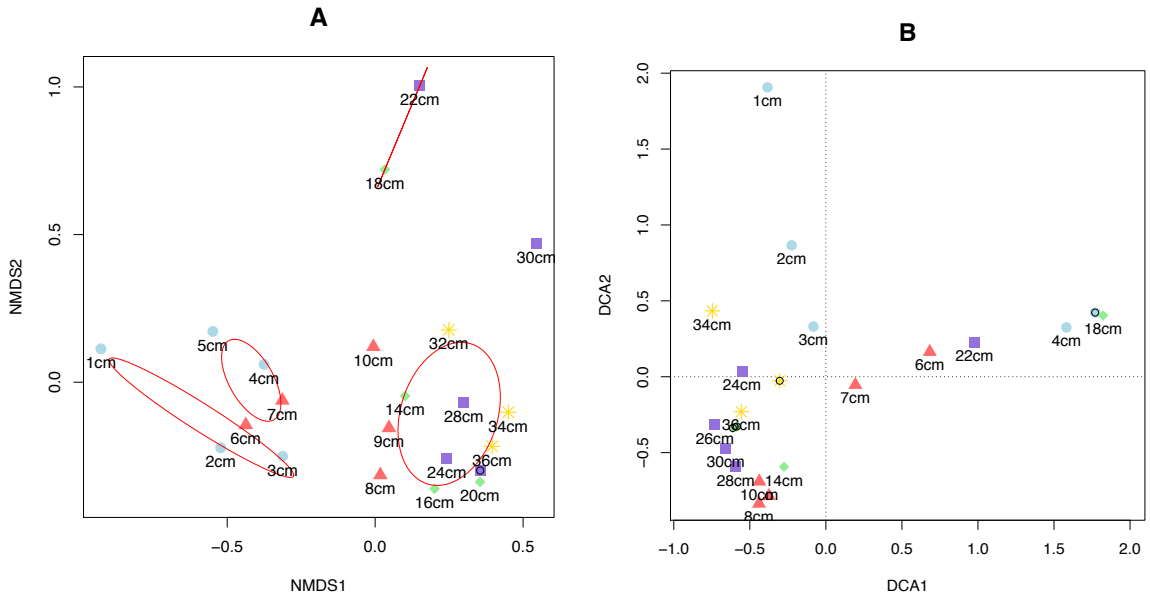


Figure 4.12. Mud Patch Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups.

4.3.5 Taxonomy Analysis

4.3.5.1 *Ravenglass Depth Core Taxonomy*

The Ravenglass depth core taxonomy derived from 16S rRNA gene sequencing analysis showed a wide range of ASVs that were more abundant in either the upper region or lower region of the core. The 10 most abundant taxa in the Ravenglass depth core are presented in a 100% stacked bar chart of relative

abundance in Figure 4.13. Individual depth profiles of 10 additional ASVs that classified to a sufficient taxonomic level to reliably discuss function are presented in Figure 4.14. AOV and Tukey HSD on abundances of these ASVs across and between sample group pairs are presented in Appendix D.8.

The Ravenglass community was almost exclusively comprised of bacterial sequences, with archaea ranging from < 0.2% of all sequences at 1 cm to 2.35% at 28 cm. The ten most abundant ASV taxonomic groups are presented in descending order of overall abundance (Figure 4.13). The most abundant ASV collapsed at genus level was the γ-*proteobacteria* genus *Woeseia* spp., and had an average abundance of 10.14 ± 0.28% between 1 – 5 cm, decreased to 5.18% at 6 cm, increased to 11.02% at 7 cm, decreased to 4.78% at 8 cm, increased to 10.91% at 9 cm, and decreased from 9.46% at 10 cm to 0.98% at 28 cm. AOV showed a statistical difference in *Woeseia* spp. abundance across all sample depth groups, and *post hoc* Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix D.8A). An uncultured *Ardenticatenales* increased from 2.81% at 1 cm to 7.94% at 5 cm, decreased to 2.71% at 10 cm, and between 12 – 28 cm, had an average abundance of 3.8 ± 0.3%. AOV and Tukey did not show statistical differences in this *Ardenticatenales* abundances (Appendix D.8B). An uncultured δ-*proteobacteria* MBNT-15 increased from 0% at 3 cm to 7.09% at 14 cm, and again to 13.51% at 26 cm. AOV showed a statistical difference in δ-*proteobacteria* MBNT-15 abundance across all sample depth groups and *post hoc* Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix D.8C). An uncultured *Anaerolineaceae* increased from 1.39% at 1 cm to 5.52% at 5 cm, and between 6 – 28 cm, had a mixed distribution with an average abundance of 3.09 ± 0.35%. AOV and Tukey did not show statistical differences in this *Anaerolineaceae* abundances (Appendix D.8D). An uncultured *Gemmatimonadetes* PAUC43f MBG increased from 0.47% at 1 cm to a peak of 5.13% at 22 cm and decreased to 4.14% at 28 cm. AOV showed a statistical difference in *Gemmatimonadetes* PAUC43f MBG abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix D.8E). An uncultured *Pirellulaceae* had an almost-identical distribution pattern as the *Gemmatimonadetes* PAUC43f MBG, increasing from 0.62% at 1 cm to 4.4% at 18 cm, decreased to 3.03% at 24 cm, and increased to 5.09% at 28 cm. AOV showed a statistical difference in *Pirellulaceae* abundance

across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, and 2 – 4, with pair 2 – 3 having a borderline significant *p* value (Appendix D.8F). An uncultured γ-*proteobacteria* decreased from 4.71% at 1 cm to 2.27% at 5 cm, increased to 4.1% at 8 cm, and decreased to 0.37% at 28 cm. AOV showed a statistical difference in γ-*proteobacteria* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix D.8G). An uncultured *Pirellulaceae* Pir4 lineage increased gradually from 1.06% at 1 cm to 4.03% at 18 cm, decreased to 2.07% at 24 cm and increased to 4.08% at 26 cm. AOV showed a statistical difference in *Pirellulaceae* Pir4 lineage abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, and 2 – 4, with pair 2 – 3 having a borderline significant *p* value (Appendix D.8H). An uncultured *Thermoanaerobaculaceae* subgroup 10 increased gradually from 0.52% at 1 cm to 5.63% at 28 cm. AOV showed a statistical difference in *Thermoanaerobaculaceae* subgroup 10 abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, and 2 – 4 (Appendix D.8I). Another uncultured *Ardenticatenales* ASV increased from 1.65% at 1 cm to 2.84% at 6 cm, decreased to 1.56% at 10 cm, and decreased gradually from 1.5% at 12 cm to 1.1% at 28 cm. AOV showed a statistical difference in *Ardenticatenales* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix D.8J).

The additional target ASV taxonomic groups are not presented in order of descending overall abundance. An uncultured *Desulfobulbaceae* had an average abundance of 4 ± 0.21% between 1 – 5 cm, decreased to 2.21% at 6 cm, and from 1.35% at 7 cm to 0.18% at 28 cm (Figure 4.14A). AOV showed a statistical difference in *Desulfobulbaceae* abundance across all sample groups and Tukey HSD showed statistical differences between all pairs of sample groups except pair 3 – 4 (Appendix D.8K). An uncultured *Chloroflexi* KD4-96 had an average abundance of 1.44 ± 0.1% between 1 – 5 cm, increased to 2% at 6 cm and between 7 – 28 cm had an average abundance of 1.96 ± 0.1% (Figure 4.14B). AOV showed a borderline-significant *p* value for *Chloroflexi* KD4-96 abundance across all sample groups, and Tukey HSD showed a statistical difference between group pair 1 – 4 (Appendix D.8L). An uncultured *Sandaracinaceae* had an average abundance of 1.56 ± 0.08% between 1

– 4m, increased to 2.59 ± 0.16% between 5 – 8 cm, and decreased to 1.22 ± 0.08% between 9 – 28 cm (Figure 4.14C). AOV and Tukey did not show statistical differences in *Sandaracinaceae* abundance across or between sample depth groups (Appendix D.8M). An uncultured *Dadabacteriales* increased gradually from 0.11% at 1 cm to 4.15% at 20 and then decreased to 2.41% at 28 cm (Figure 4.14D). AOV showed a statistical difference in *Dadabacteriales* abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix D.8N). A second uncultured *Desulfobulbaceae* had an average abundance of 3.47 ± 0.58% between 1 – 4 cm, 1.71 ± 0.44% between 5 – 8 cm, and 0.17 ± 0.04% between 9 – 28 cm (Figure 4.14E). AOV showed a statistical difference in the second *Desulfobulbaceae* abundance across all sample groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix D.8O). An uncultured *Saprospiraceae* decreased from 3.65% at 1 cm to 0.16% at 16 cm, and to 0% at 28 cm (Figure 4.14F). AOV showed a statistical difference in *Saprospiraceae* abundance across all sample depth groups and Tukey HSD showed statistical differences between all pairs of groups except 3 – 4 (Appendix D.8P). An uncultured *Prolixibacteraceae* increased from 0.95% at 1 cm to 6.12% at 4 cm and decreased to 0% at 10 cm and remained at 0% throughout (Figure 4.14G). AOV showed a statistical difference in *Prolixibacteraceae* abundance across all sample groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix D.8Q). The *Methyloceanibacter* spp. increased gradually from 0.42% at 1 cm to 2.2% at 28 cm (Figure 4.14H). AOV showed a statistical difference in *Methyloceanibacter* spp. abundance across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix D.8R). The *Thaumarchaeota* genus *Nitrosopumilus* spp. had an average abundance of 0.07 ± 0.04% between 1 – 4 cm, increased from 0.47% at 5 cm to 1.06% at 9 cm, decreased to 0.22% at 12 cm, and increased to 1.03% at 28 cm (Figure 4.14I). AOV showed a statistical difference in *Nitrosopumilus* spp. abundance across all sample groups and Tukey HSD showed a statistical difference between group pair 1 – 4, and borderline-significant *p* value between pair 1 – 2 (Appendix D.8S). An uncultured *Bathyarchaeia* decreased from 0.06% at 1 cm to 0% at 3 cm, increased from 0.18% at 4 cm to 0.66% at 8 cm, and between 9 – 28 cm had an average abundance of 0.36 ± 0.05% (Figure 4.14J). AOV and Tukey did not show statistical differences in *Bathyarchaeia* abundance across or between sample groups (Appendix D.8T).

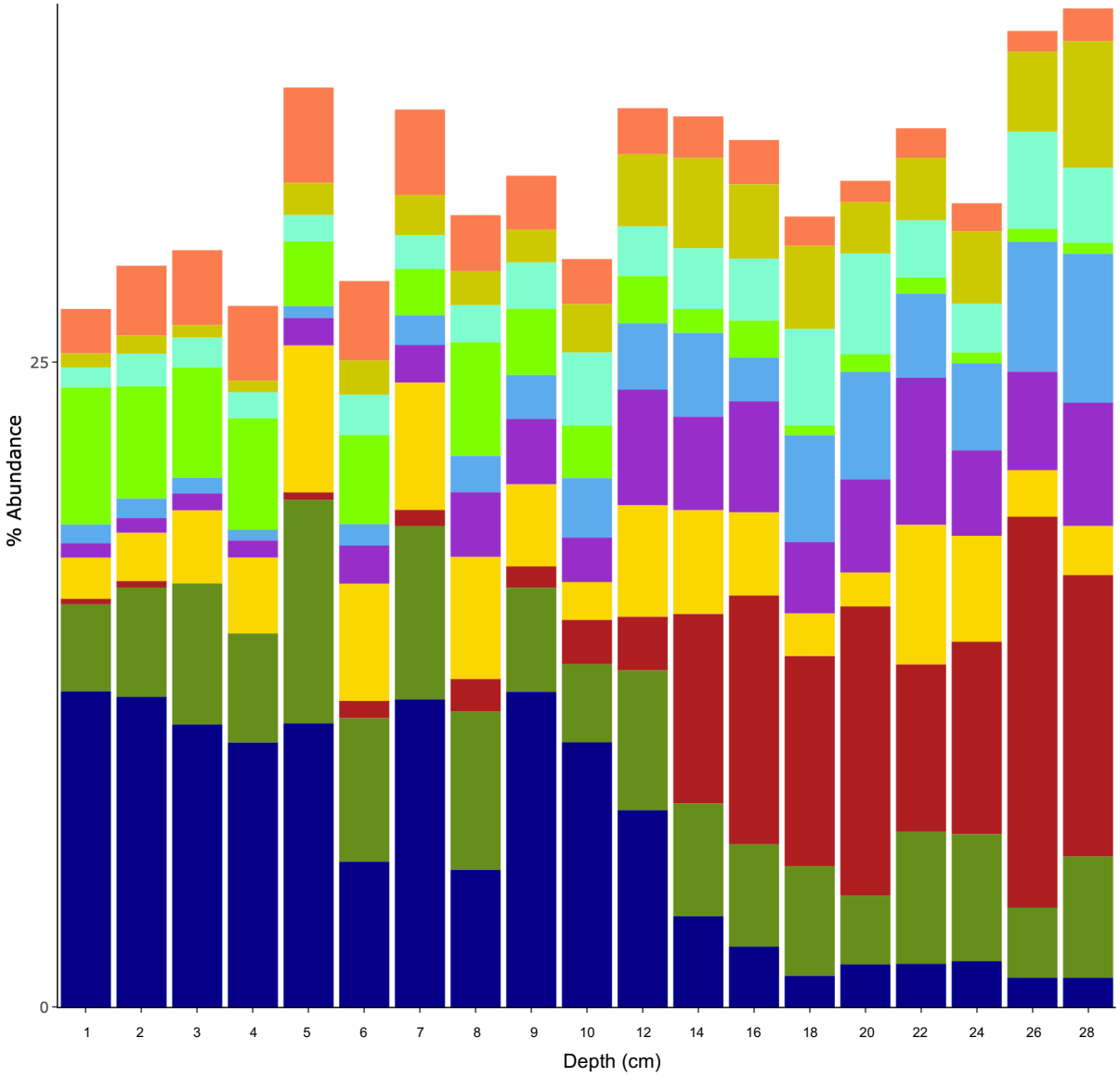


Figure 4.13. Ravenglass Depth Core: 100% stacked bar chart of ASV taxonomy collapsed at genus level, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = *Woeseia* spp.; dark green = uncultured *Ardenticatenales* (1); red = uncultured δ-proteobacteria MBNT-15; yellow = uncultured *Anaerolineaceae*; purple = uncultured *Gemmatimonadetes* PAUC43f marine benthic group; blue = uncultured *Pirellulaceae*; green = uncultured γ-proteobacteria; light blue = *Pirellulaceae* Pir4 lineage; gold = *Thermoanaerobaculaceae* subgroup 10; orange = uncultured *Ardenticatenales* (2).

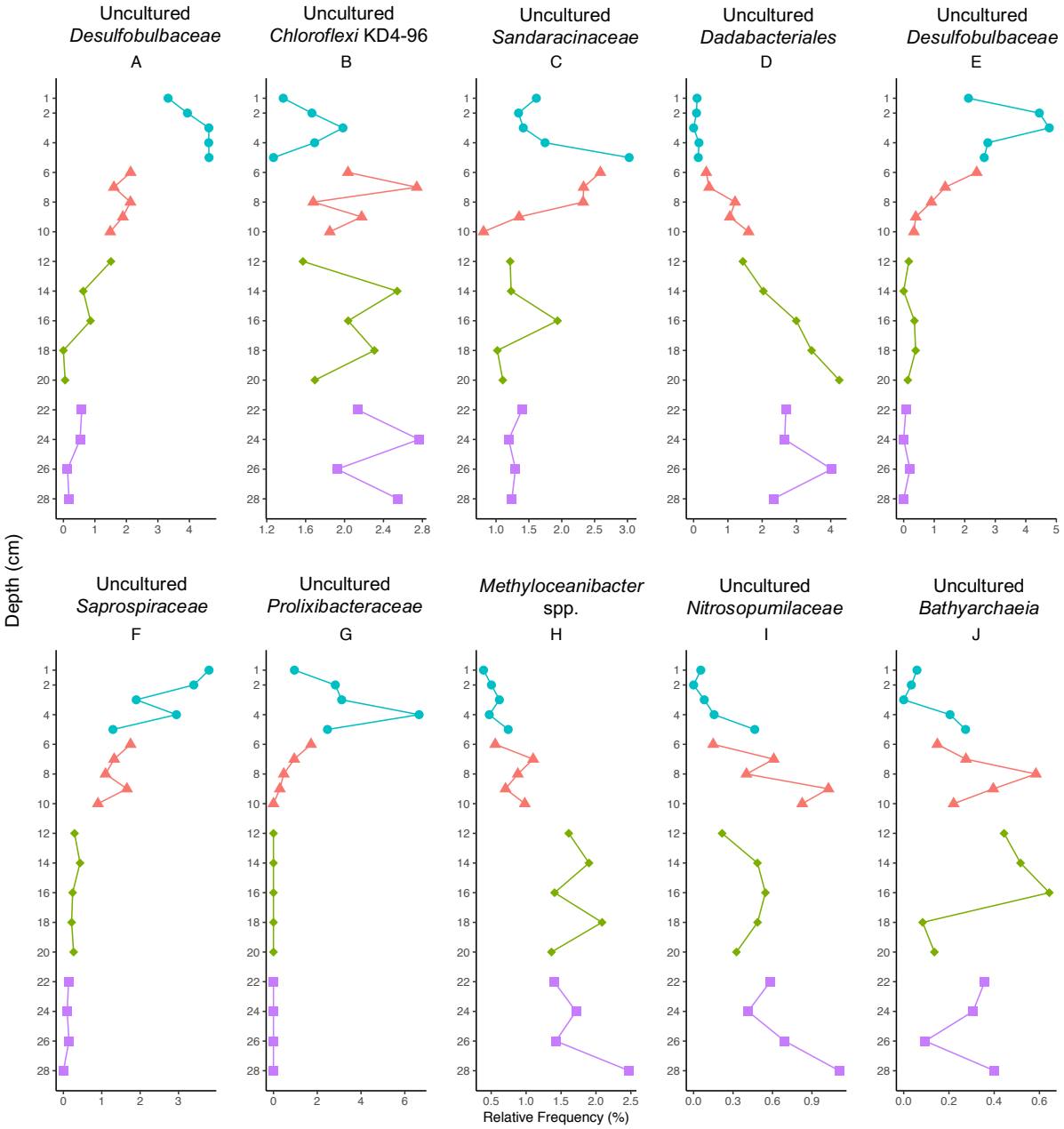


Figure 4.14. Ravenglass Depth Core: Depth profiles of ten additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). Uncultured *Desulfobulbaceae (1)*; (B). Uncultured *Chloroflexi* KD4-96; (C). Uncultured *Sandaracinaceae*; (D). Uncultured *Dadabacteriales*; (E). Uncultured *Desulfobulbaceae (2)*; (F). Uncultured *Saprospiraceae*; (G). Uncultured *Prolixibacteraceae*; (H). *Methyloceanibacter* spp.; (I). Uncultured *Nitrosopumilaceae*; (J). Uncultured *Bathyarchaeia*.

4.3.5.2 *Mud Patch Depth Core Taxonomy*

The microbial community within the Mud Patch depth core was more homogenous than that observed at Ravenglass as indicated in the nMDS and DCA plots (Figure 4.8 and 4.12), with the core characterised by some discrete depth samples being comprised of a very high percentage of a single or only a few ASVs.

The 10 most abundant taxa in the Mud Patch depth core are presented in a 100% stacked bar chart of relative abundance in Figure 4.15. Individual depth profiles of six additional ASVs that classified to a sufficient taxonomic level to reliably discuss function are presented in Figure 4.16. AOV and Tukey HSD on abundances of these ASVs across and between sample group pairs are presented in Appendix D.9. Similar to the Ravenglass community, the Mud Patch was almost exclusively comprised of bacterial sequences, with archaea ranging from < 0.01% of all sequences at 5 cm to 2.98% at 36 cm.

The ten most abundant ASV taxonomic groups are presented in descending order of overall abundance (Figure 4.15). The most abundant ASV collapsed at genus level was an uncultured *Planococcaceae* which had a very sporadic distribution, with abundances of 53.5% at 4 cm, 49% at 5 cm, 12% at 6 cm, 63.5% at 18 cm, and 26.5% at 22 cm. The average abundance of this ASV excluding these samples was 0.5 ± 0.2%. AOV and Tukey HSD did not show a statistical difference in *Planococcaceae* abundance across or between sample groups (Appendix D.9A). An uncultured *Desulfobulbaceae* increased from 1.99% at 1 cm to 11.44% at 10 cm, decreased to 1.56% at 18 cm, and between 20 – 36 cm had an average abundance of 3.65 ± 0.27%. AOV showed a statistical difference in *Desulfobulbaceae* abundance across all sample depth groups and Tukey HSD showed statistical differences between sample group 2 and all other groups (Appendix D.9B). An uncultured *Flavobacteriaceae* peaked at 17% at 2 cm, decreased to 1.37% at 5 cm, and between 6 – 36 cm had an average abundance of 2.27 ± 0.29%. AOV showed a statistical difference in *Flavobacteriaceae* abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 4, with borderline significant differences between 1 – 3 and 1 – 5 (Appendix D.9C). An uncultured γ-*proteobacteria* B2M28 increased from 1.03% at 1 cm to 7.64% at 14 cm, decreased to 1.17% at 18 cm, and between 20 – 36 cm had an average abundance of 3.98 ± 0.24%. AOV showed a statistical difference in γ-*proteobacteria* B2M28 abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2, 1 – 3, and borderline significance between 1 – 4 and 1 – 5 (Appendix D.9D). The *Psychrobacter* spp. was 39.62% at 1 cm, 6.52% at 4 cm, 9.22% at 5 cm, 4.36% at 22 cm, 3.9% at 24 cm, 8.41% at 32 cm, and 9.19% at 34 cm, with an average abundance without these samples of 0.12 ± 0.06%. AOV and Tukey did not show statistical differences in the *Psychrobacter* spp. abundance

(Appendix D.9E). An uncultured γ-*proteobacteria* increased from 1.19% at 1 cm to 4.58% at 3 cm, decreased to 1.86% at 5 cm, increased to 5.17% at 6 cm, decreased to 0.77% at 18 cm, and between 20 – 36 cm had an average abundance of 3 ± 0.16%. AOV showed a statistical difference in γ-*proteobacteria* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 1 – 4, and borderline significance between 2 – 3 and 2 – 5 (Appendix D.9F). An uncultured *Syntrophobacteraceae* increased from 0.7% at 1 cm to 6.7% at 28 cm and decreased to 2.7% at 36 cm. AOV showed a statistical difference in *Syntrophobacteraceae* abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 4 and 1 – 5 (Appendix D.9G). An uncultured *Anaerolineaceae* increased from 0.61% at 1 cm to 3.07% at 16 cm, then was 0.81%, 3.88%, and 18.1% at 18, 20, and 22 cm respectively, and between 24 – 36 cm had an average abundance of 3.51 ± 0.15%. AOV showed a statistical difference in *Anaerolineaceae* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 1 – 5, 2 – 5, and borderline significance between 1 – 2 (Appendix D.9H). A second uncultured γ-*proteobacteria* had a more sporadic distribution with depth, ranging from 0.8% at 1 cm to 5.43% at 7 cm, and between 8 – 36 cm had an average abundance of 2.07 ± 0.16%. AOV and Tukey did not show statistical differences in this second γ-*proteobacteria* across or between sample depth groups (Appendix D.9I). A second uncultured *Desulfobulbaceae* had an average abundance of 0.34 ± 0.06% between 1 – 5 cm, increased from 1.33% at 6 cm to 2.99% at 9 cm, and after 16 cm had a sporadic distribution ranging from 0.59% at 18 cm to 3.55% at 30 cm. AOV showed a statistical difference in this second *Desulfobulbaceae* across all sample groups and Tukey HSD showed statistical differences between sample group 1 and all other groups (Appendix D.9J).

The additional target ASV taxonomic groups are not presented in order of descending overall abundance. The *Lutimonas* spp. was 5.07%, 7.44%, and 7.78% at 1, 2, and 3 cm respectively, and between 4 – 36 cm had an average abundance of 0.78 ± 0.06% (Figure 4.16A). AOV showed a statistical difference in *Lutimonas* spp. abundance across all sample depth groups and Tukey HSD showed a statistical difference between group pair 1 – 4, and borderline significances between group pairs 1 – 2, 1 – 3, and 1 – 5 (Appendix D.9K). The *Woeseia* spp. had a very mixed abundance with depth, ranging from 0% at 18 cm and 2.98% at 14 cm, with an

average abundance of 1.66 ± 0.17% across the whole core (Figure 4.16B). AOV and Tukey did not show statistical differences in the *Woeseia* spp. abundance across or between sample depth groups (Appendix D.9L). The *Thermoanaerobaculaceae* subgroup 10 increased gradually from 0.63% at 1 cm to 2.14% at 16 cm, decreased to 0.51% at 18 cm, and between 20 – 36 cm had an average abundance of 1.96 ± 0.14% (Figure 4.16C). AOV showed a statistical difference in *Thermoanaerobaculaceae* subgroup 10 abundance across all sample depth groups and Tukey HSD showed a statistical difference between group pair 1 – 4, and borderline significance between pairs 1 – 2 and 1 – 5 (Appendix D.9M). The *Methyloceanibacter* spp. increased gradually with depth from 0.8% at 1 cm to 2.14% at 32 cm (Figure 4.16D). AOV showed a statistical difference in *Methyloceanibacter* spp. abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 5 and 2 – 5, and borderline significance between 1 – 4 and 3 – 5 (Appendix D.9N). An uncultured δ-*proteobacteria* SEEP-SRB1 increased from 0.15% at 1 cm to 2.72% at 30 cm (Figure 4.16E). AOV showed a statistical difference in δ-*proteobacteria* SEEP-SRB1 abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 1 – 5, 2 – 4, and 2 – 5 (Appendix D.9O). An uncultured archaeal *Bathyarchaeia* had an average abundance of 0.04 ± 0.01% between 1 – 10 cm and increased from 0.23% at 14 cm to 1.25% at 36 cm (Figure 4.16F). AOV showed a statistical difference in *Bathyarchaeia* abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 5 and 2 – 5, and borderline significance between 3 – 5 (Appendix D.9P).

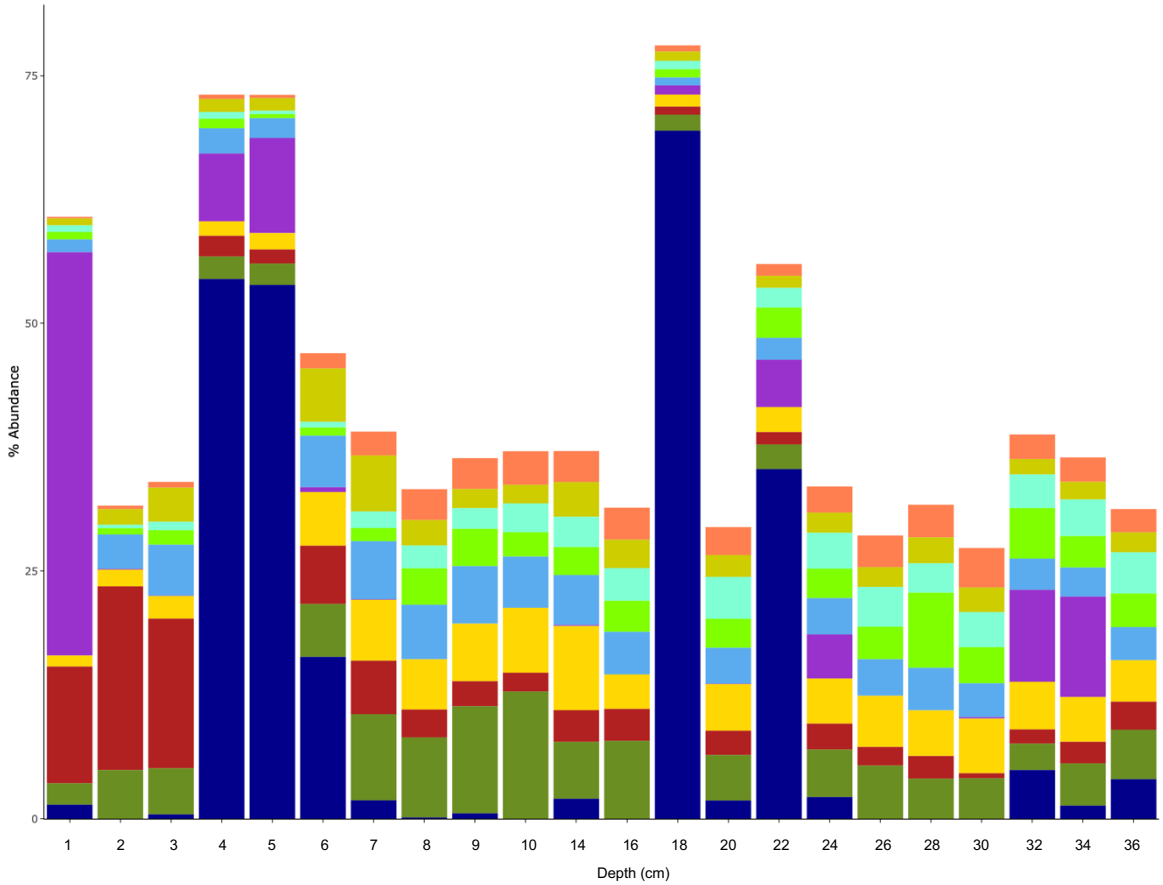


Figure 4.15. Mud Patch Depth Core: 100% stacked bar chart of ASV taxonomy collapsed at genus level, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = uncultured *Planococcaceae*; dark green = uncultured *Desulfobulbaceae* (1); red = uncultured *Flavobacteriaceae*; yellow = uncultured γ-*proteobacteria* B2M28; purple = *Psychrobacter* spp.; blue = uncultured γ-*proteobacteria* (1); light green = uncultured *Syntrophobacteraceae*; light blue = uncultured *Anaerolineaceae*; gold = uncultured γ-*proteobacteria* (2); orange = uncultured *Desulfobulbaceae* (2).

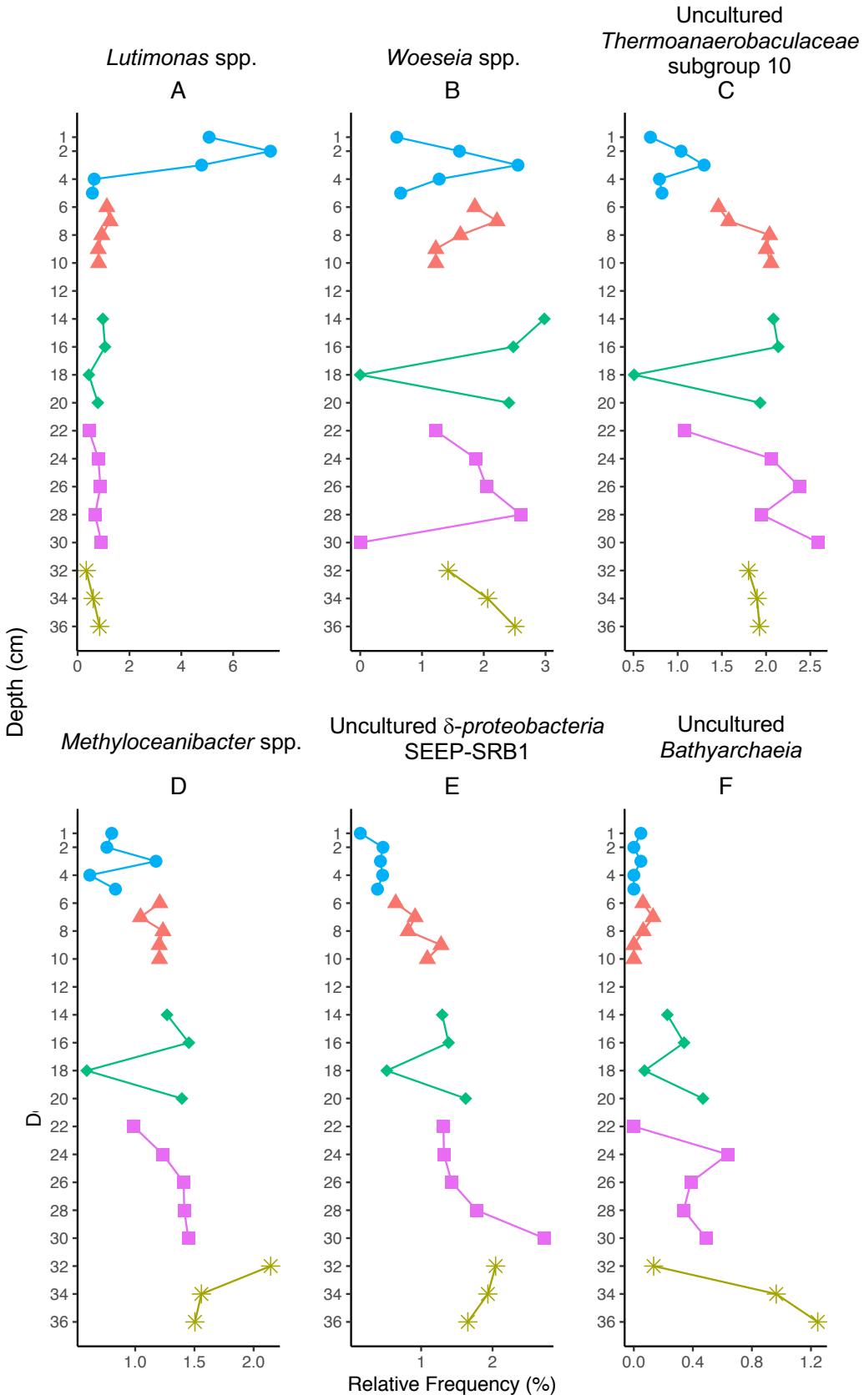


Figure 4.16. Mud Patch Depth Core: Depth profiles of ten additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). *Lutimonas* spp.; (B). *Woeseia* spp.; (C). Uncultured *Thermoanaerobaculaceae* subgroup 10; (D). *Methyloceanibacter* spp.; (E). Uncultured δ-*proteobacteria* SEEP-SRB1; (F). Uncultured *Bathyarchaeia*.

4.3.6 Tax4Fun Inferred Functional Gene Abundance Profiles

A total of 6,455 genes were predicted by Tax4Fun in the Ravenglass core, and 6,555 genes in the Mud Patch core. Target gene abundances reflect the percentage normalised abundance of the gene at a given sample depth.

4.3.6.1 *Ravenglass Depth Core Inferred Functional Gene Abundance Profile*

The five functional genes targeted during end point PCR were identified in the output functional gene abundance profile inferred by Tax4Fun from the Ravenglass depth core. The *amoA* gene (not distinguished between archaeal and bacterial variants of the gene) (Figure 4.17A) increased between 1 – 8 cm from 0.12% to 0.57% (average = 0.37 ± 0.1%), increased to 1.8% at 9 cm, and between 10 – 28 cm, averaged 0.9 ± 0.1%. AOV of the *amoA* gene abundance was significantly different across all sample groups, with Tukey HSD showing a statistical difference between group pair 1 – 4, and a nearly significant difference between 1 – 2 (*p* value of 0.054) (Appendix D.10A). The denitrifying *nirK* gene (Figure 4.17B) followed an identical pattern to *amoA* albeit with higher abundances: 2.1 ± 0.4% between 1 – 8 cm; 6.6% at 9 cm, and; 3.9 ± 0.4% between 10 – 28 cm. AOV of the *nirK* gene abundance showed a statistical difference in gene abundance across all sample groups, with Tukey HSD showing a statistical difference between sample group pairs 1 – 2 and 1 – 4 (Appendix D.10B). The other nitrite reductase, *nirS* (Figure 4.17C), similarly increased with depth, though with lower abundances than the *nirK* gene: 1.2 ± 0.1% between 1 – 7 cm; 2.4% at 8 cm, and; 2.6 ± 0.3% between 9 – 28 cm. AOV of the *nirS* gene abundance showed a statistical difference in abundance across all sample groups, with Tukey HSD showing a statistical difference between group pairs 1 – 3 and 1 – 4 (Appendix D.10C). The *dsrA* gene (Figure 4.17D) was most abundant in the upper region of the core: between 1 – 6 cm, the average abundance was 2.1 ± 0.2%, and dropped to 0.4 ± 0.1% between 7 – 28 cm, with AOV showing a statistical difference in abundance across all groups and Tukey showing a statistical difference between group 1 and all other groups, and pair 2 – 4 (Appendix D.10D). Negligible amounts of *mcrA* (*i.e.*, lower than 0.0003%) was inferred by Tax4Fun (Figure 4.17E), and AOV only showed borderline significance in *mcrA* abundance (Appendix D.10E). Scatterplots and Spearman correlations between the abundances of the same functional gene as inferred from end point PCR and by Tax4Fun showed no correlation between archaeal *amoA* (Figure 4.18A), bacterial *amoA* (Figure 4.18B),

and *nirS* (Figure 4.18D) abundances, but showed a strong positive correlation between *dsrA* abundances (Figure 4.18E; $\rho = 0.92$, p-value = < 0.005) and a negative correlation for *nirK* (Figure 4.18C).

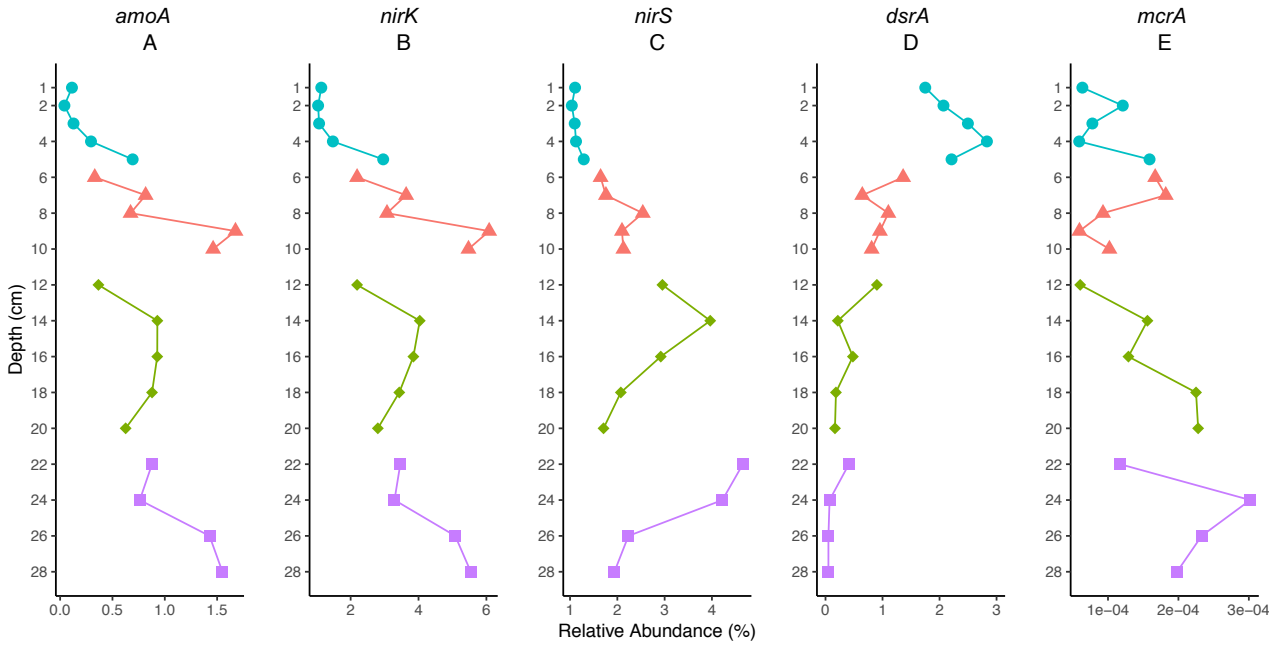


Figure 4.17. Ravenglass Depth Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA*.

Nitrogen cycling genes were generally more abundant in the upper region of the Ravenglass core. The nitrate reductase (NO_3^{2-} to NO_2^-) *narG* increased gradually from 0.47% at 1 cm to 5.42% at 7 cm and decreased gradually from 4.04% at 8 cm to 2.86% at 28 cm (Figure 4.19A). AOV showed a statistical difference in *narG* abundance across all sample depth groups and Tukey HSD showed statistical differences in abundance between group pairs 2 – 3 and 2 – 4, and borderline significance between 1 – 4 (Appendix D.11A). The other nitrate reductase *napA* increased from 2.33% at 1 cm to 2.66% at 4 cm, decreased to 1.17% at 7 cm, increased to 1.9% at 8 cm, and decreased to 0.83% at 28 cm (Figure 4.19B). AOV showed a statistical difference in *napA* abundance across all sample depth groups and Tukey HSD showed statistical

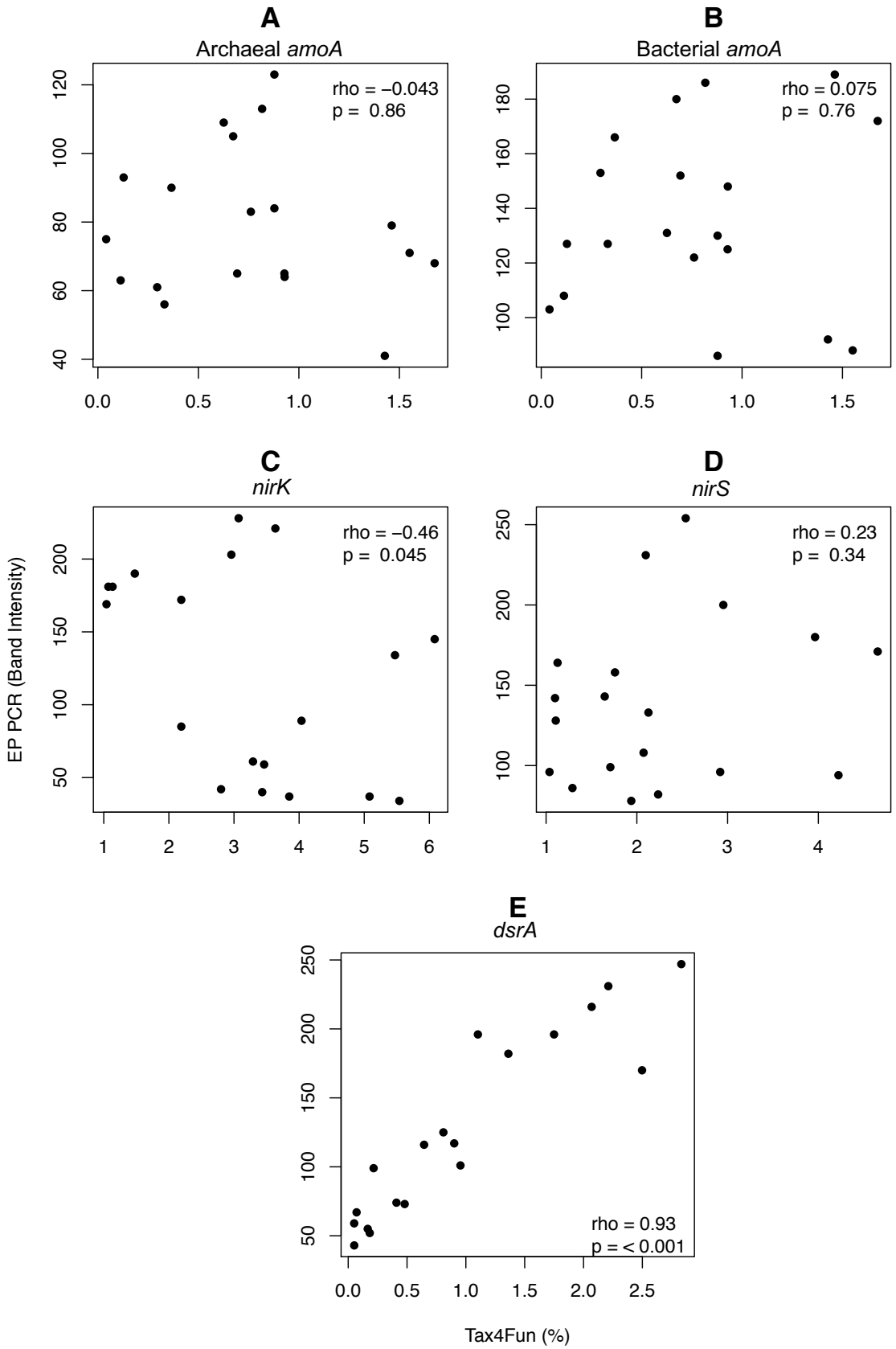


Figure 4.18. Ravenglass Depth Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA*.

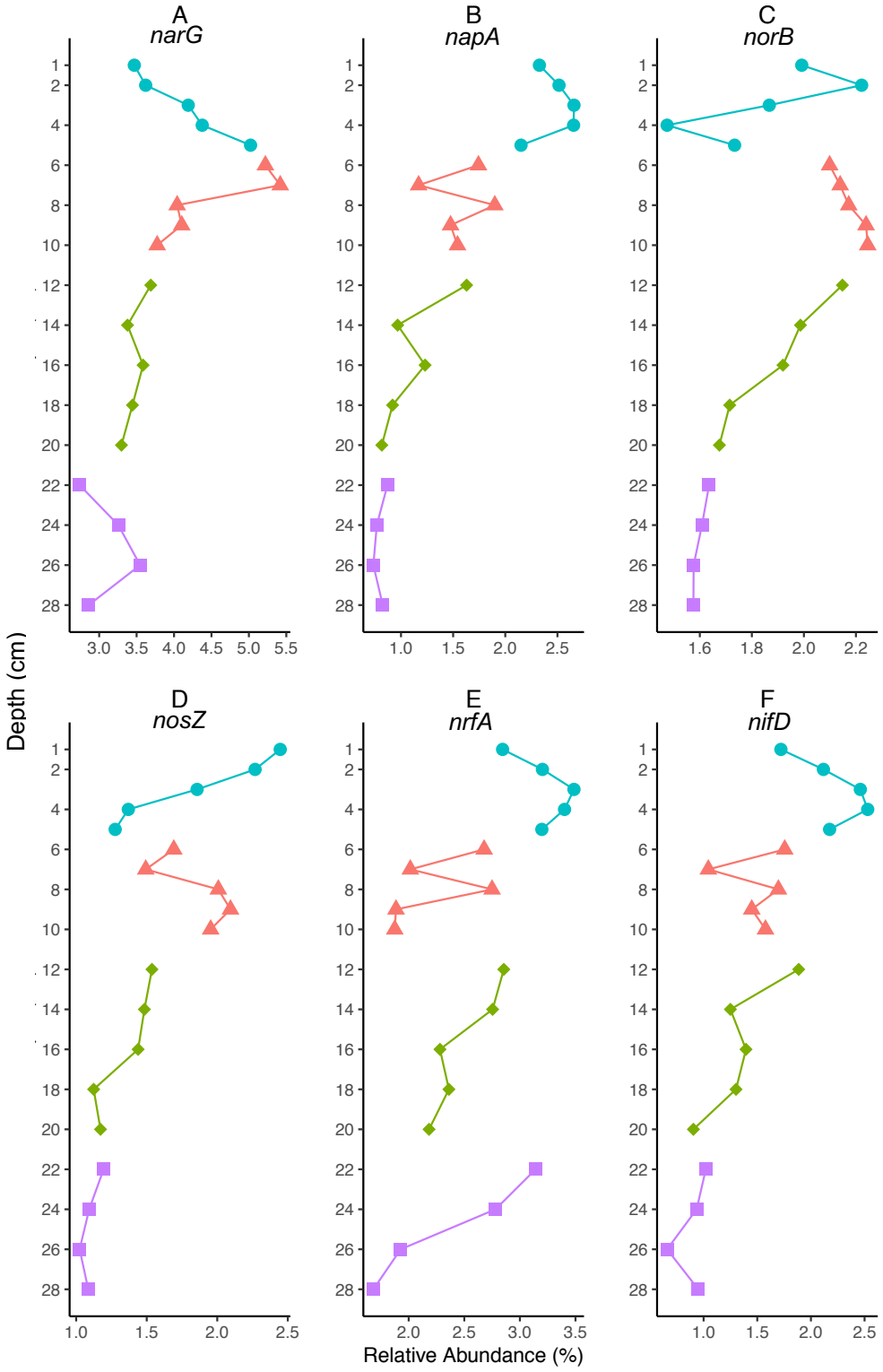


Figure 4.19. Ravenglass Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD*.

differences between group 1 and all other groups, between 2 – 4, and borderline significance between 2 – 3 (Appendix D.11B). The nitric oxide reductase (NO to N_2O)

norB decreased from 2.22% at 2 cm to 1.47% at 4 cm, increased from 2.1% at 6 cm to 2.25% at 10 cm, and decreased to 1.58% at 28 cm (Figure 4.19C). AOV showed a statistical difference in *norB* abundance across all sample depth groups and Tukey HSD showed a statistical difference between group pair 2 – 4, and borderline significant differences between group pairs 1 – 2 and 2 – 3 (Appendix D.11C). The nitrous-oxide reductase (N_2O to N_2) *nosZ* decreased from 2.45% at 1 cm to 1.28% at 5 cm, increased from 1.69% at 6 cm to 2.09% at 9 cm, and decreased gradually to 1.09% at 28 cm (Figure 4.19D). AOV showed a statistical difference in *nosZ* abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 4 and 2 – 4 (Appendix D.11D). The dissimilatory nitrate reduction (DNRA) pathway nitrite reductase (HNO_2 to NH_3) *nrfA* increased from 2.84% at 1 cm to 3.49% at 3 cm, decreased to 2.02% at 7 cm, and between 8 – 28 cm had an average abundance of 2.37 ± 0.17% (Figure 4.19E). AOV showed a statistical difference in *nrfA* abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 1 – 4, and borderline significant difference between 1 – 3 (Appendix D.11E). The nitrogen fixating nitrogenase (N_2 to NH_3) *nifD* gene increased from 1.72% at 1 cm to 2.53% at 4 cm, decreased to 1.04% at 7 cm, and between 8 and 28 cm had an average abundance of 1.25 ± 0.11% (Figure 4.19F). AOV showed a statistical difference in *nifD* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and between pair 2 – 4 (Appendix D.11F).

Other target genes included a mix of aerobic and anaerobic functional genes and are not presented in a specific order. The Wood-Ljungdahl anaerobic carbon fixing gene *acsB* increased from 1.64% at 1 cm to 2.67% at 4 cm, decreased from 2.29% at 5 cm to 0.56% at 7 cm, and decreased to 0.02% at 28 cm (Figure 4.20A). AOV showed a statistical difference in *acsB* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and borderline significant difference between group pair 2 – 4 (Appendix D.12A). The dissimilatory sulfate reduction pathway gene *aprA* increased from, 2.63% at 1 cm to 4.29% at 4 cm, and decreased from 3.33% at 5 cm to 0.12% at 28 cm (Figure 4.20B). AOV showed a statistical difference in *aprA* abundance across all sample depth groups and Tukey HSD showed statistical differences between all pairs of sample groups except 3 – 4 (Appendix D.12B). The aerobic methane

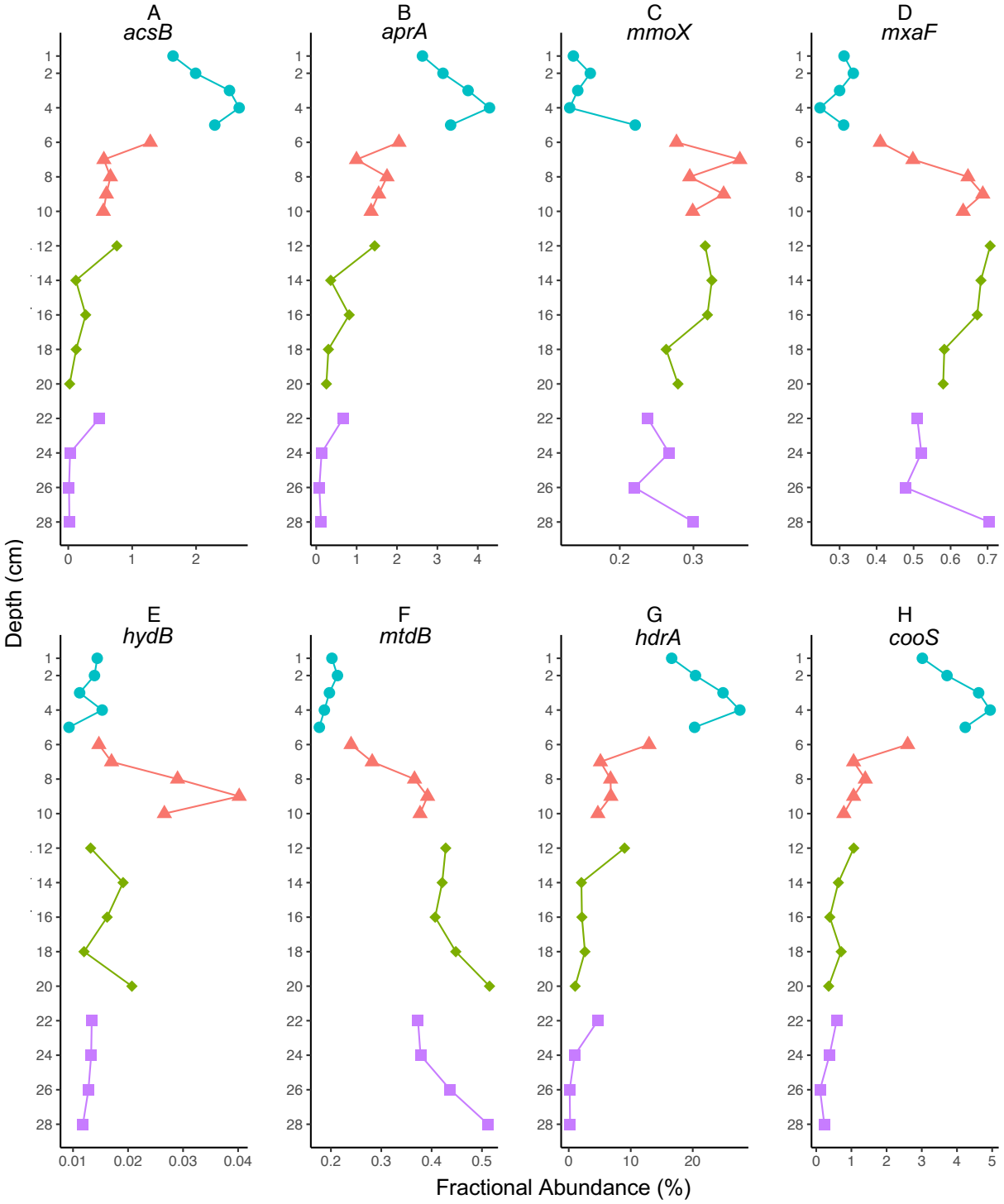


Figure 4.20. Ravenglass Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS*.

monooxygenase *mmoX* from the methane oxidation pathway had an average abundance of 0.14 ± 0.01% between 1 – 4 cm, increased from 0.22% at 5 cm to 0.36% at 7 cm, and between 8 – 28 cm had an average abundance of 0.29 ± 0.01% (Figure 4.20C). AOV showed a statistical difference in *mmoX* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1

and all other groups, and borderline significant difference between 2 – 4 (Appendix D.12C). The other methane oxidation gene, methanol dehydrogenase *mxaF* had an average abundance of 0.3 ± 0.02% between 1 – 4 cm, increased from 0.25% at 5 cm to 0.65% at 8 cm, and between 9 – 28 cm had an average abundance of 0.61 ± 0.03% (Figure 4.20D). AOV showed a statistical difference in *mxaF* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix D.12D). The NiFe hydrogenase *hydB* had a low abundance throughout, not exceeding 0.040% at 9 cm (Figure 4.20E). AOV showed a statistical difference in *hydB* abundance across all sample depth groups with Tukey showing statistical differences between group pairs 1 – 2 and 2 – 4, and borderline significant difference between 2 – 3 (Appendix D.12E). The methanotrophic *mtdB* gene had an average abundance of 0.2 ± 0.01% between 1 – 5 cm and increased gradually to 0.52% at 20 cm (Figure 4.20F). AOV showed a statistical difference in *mtdB* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other pairs, between 2 – 3, and borderline significant difference between group pair 2 – 4 (Appendix D.12F). The heterodisulfide reductase *hdrA* is linked only to methanogenesis according to the KEGG Ortholog database but has recently been linked to anaerobic methane oxidation (Wang *et al.*, 2014), and increased from 16.6% at 1 cm to 27.6% at 4 cm and decreased from 20.3% at 5 cm to 0.22% at 28 cm (Figure 4.20G). AOV showed a statistical difference in *hdrA* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and borderline significant difference between pair 2 – 4 (Appendix D.12G). Another anaerobic carbon fixation gene *cooS* increased from 3.02% at 1 cm to 4.94% at 5 cm, and decreased from 4.23% at 6 cm to 0.21% at 28 cm (Figure 4.20H). AOV showed a statistical difference in *cooS* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and borderline significant difference between pair 2 – 4 (Appendix D.12H).

4.3.6.2 *Mud Patch Depth Core Inferred Functional Gene Abundance Profile*

The five functional genes targeted during end point PCR were identified in the output functional gene abundance profile inferred by Tax4Fun from the Mud Patch depth core. The *amoA* gene (Figure 4.21A) had a sporadic distribution with depth, with no discernible pattern. The average abundance throughout the core was 0.1 ±

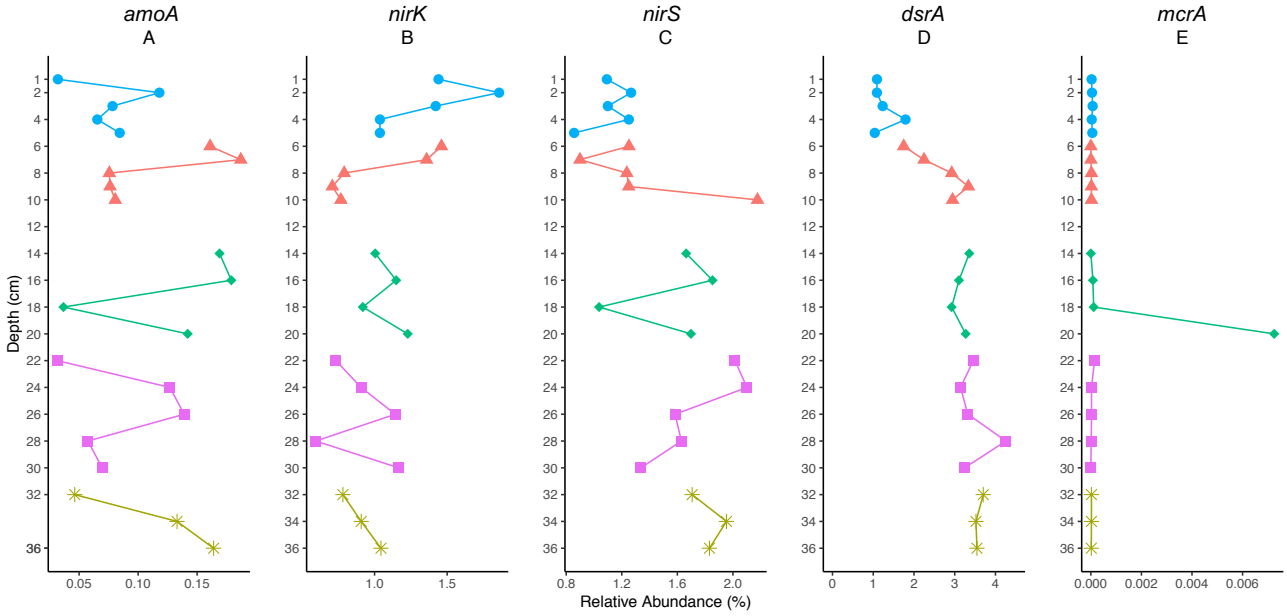


Figure 4.21. Mud Patch Depth Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA*.

0.01% and peaked at 7 cm with 0.22% and a low of 0.03% at 1 and 22 cm. AOV of the *amoA* inferred gene abundance did not show a statistical difference between sample groups (Appendix D.13A). The *nirK* gene (Figure 4.21B) also had sporadic distribution with depth, although was generally higher in abundance between 1 – 7 cm (average = 1.4 ± 0.1%) than between 8 – 36 cm (1 ± 0.06%). AOV of the *nirK* inferred gene abundance did not show a statistical difference between sample groups (Appendix D.13B). By contrast, the *nirS* gene (Figure 4.21C) was lower in abundance in the upper part of the core, with an average abundance of 1.1 ± 0.06% between 1 – 7 cm, and 1.82 ± 0.1% between 8 – 36 cm. AOV of the *nirS* gene abundance showed a statistical difference across sample groups and Tukey HSD showed a statistical difference between sample group pairs 1 – 4 and 1 – 5 (Appendix D.13C). The *dsrA* gene (Figure 4.21D) had a similar distribution to the *nirS*, increasing in abundance from 0.8% at 1 cm to 1.9% at 7 cm (average = 1.23 ± 0.2%), and with an average abundance between 8 – 36 cm of 3 ± 0.1%. AOV showed a statistical difference in *dsrA* inferred gene abundance and Tukey HSD showed a statistical difference between group 1 and all other groups (Appendix D.13D). Negligible amounts of *mcrA* gene was predicted by Tax4Fun (Figure 4.21E, Appendix D.13E). Scatterplots and Spearman correlations between the abundances of the same functional gene as inferred from end point PCR and by Tax4Fun showed

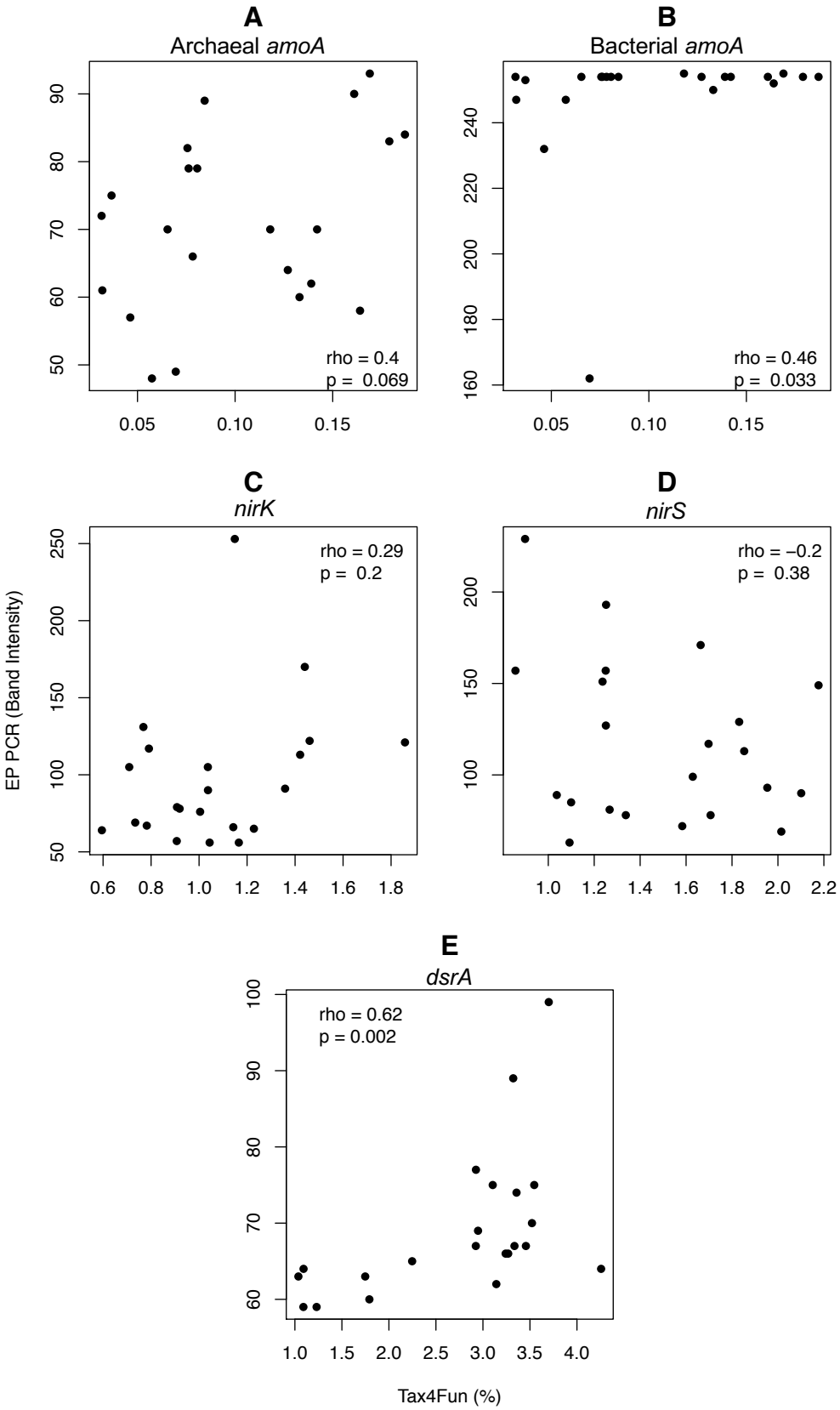


Figure 4.22. Mud Patch Depth Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) and the abundance of the same gene as inferred by Tax4Fun (relative abundance). Spearman correlation is represented by ρ ('rho') and p values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*; (E). *dsrA*.

no correlation between archaeal *amoA* (Figure 4.22A), *nirK* (Figure 4.22C), and *nirS* (Figure 4.22D) abundances, but showed a strong positive correlation between bacterial *amoA* (Figure 4.22B) *dsrA* abundances (Figure 4.22D).

Of the nitrogen cycling genes, the *narG* gene gradually increased from 4.72% at 1 cm to a peak of 7.23% at 28 cm (Figure 4.23A). AOV showed a statistical difference in *narG* abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, and 1 – 5 (Appendix D.14A). The *napA* gene increased from 1.98% at 2 cm to 4% at 28 cm (Figure 4.23B). AOV showed only borderline significant difference in *napA* abundance across all sample depth groups (Appendix D.14B). The *norB* gene increased from 2.89% at 3 cm to 4.44% at 28 cm (Figure 4.23C). AOV showed only borderline significant difference in *norB* abundance across all sample depth groups (Appendix D.14C). The *nosZ* gene peaked at 3.05% at 2 cm and between 4 – 36 cm had an average abundance of 2.16 ± 0.05% (Figure 4.23D). AOV and Tukey did not show statistical differences in *nosZ* across or between sample depth groups (Appendix D.14D). The *nrfA* gene increased progressively from 1.55% at 1 cm to 3.68% at 10 cm, and between 14 – 36 cm had an average abundance of 3.08 ± 0.1% (Figure 4.23E). AOV showed a statistical difference in *nrfA* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix D.14E). The *nifD* gene had an average abundance of 2.03 ± 0.13% between 1 – 6 cm, and 3.8 ± 0.1% between 7 – 36 cm (Figure 4.23F). AOV showed a statistical difference in *nifD* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and borderline significant difference between group pair 2 – 4 (Appendix D.14F).

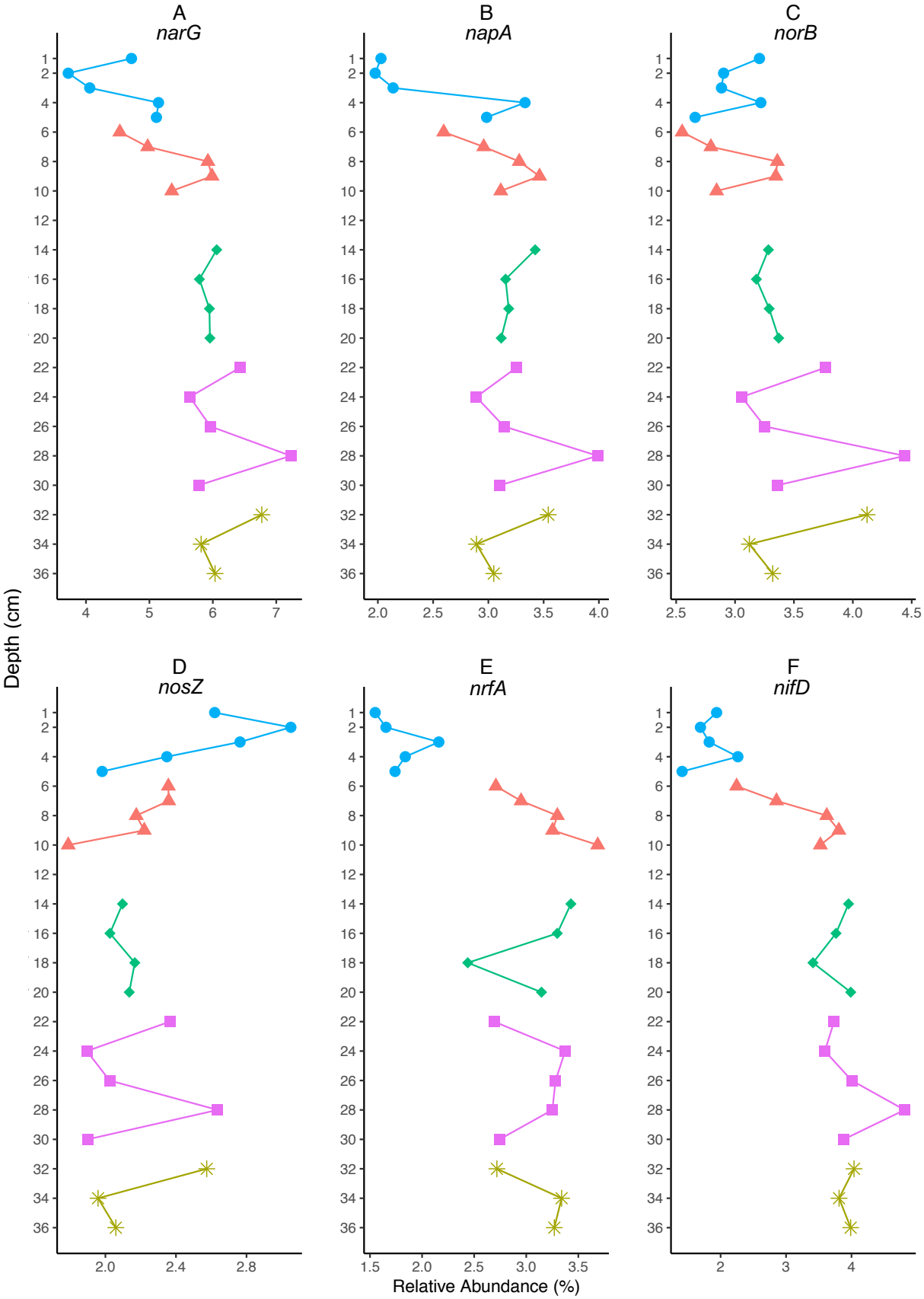


Figure 4.23. Mud Patch Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD*.

The *acsB* gene increased from an average abundance between 1 – 7 cm of 1.85 ± 0.16% to an average abundance between 8 – 36 cm of 4.46 ± 0.19% (Figure 4.24A). AOV showed a statistical difference in *acsB* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and between pairs 2 – 4 and 2 – 5 (Appendix D.15A). The *aprA* gene increased from an average abundance of 2.25 ± 0.21% between 1 – 7 to 4.92 ± 0.15% between 8 – 36 cm (Figure 4.24B). AOV showed a statistical difference in *aprA* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and between pairs 2 – 4 and 2 – 5 (Appendix D.15B). The *mmoX* gene decreased from 0.31% at 1 cm to 0.17% at 5 cm and between 6 – 36 cm had an average abundance of 0.21 ± 0.13% (Figure 4.24C). AOV and Tukey did not show statistical differences in *mmoX* abundance across or between sample depth groups (Appendix D.15C). The *mxaF* gene had an average abundance of 0.55 ± 0.03% across the core, including a peak to 1.01% at 30 cm (Figure 4.24D). AOV and Tukey did not show statistical differences in *mxaF* abundance across or between sample depth groups (Appendix D.15D). The *hydB* gene had an average abundance of 0.017 ± 0.001% across the whole core (Figure 4.24E). AOV showed a statistical difference in *hydB* abundance across all sample depth groups and Tukey HSD showed statistical differences between group pairs 1 – 4 and 1 – 5 (Appendix D.15E). The *mtdB* gene had an average abundance of 0.29 ± 0.01% across the whole core including a peak of 0.47% at 30 cm (Figure 4.24F). AOV and Tukey did not show statistical differences in *mtdB* abundance across or between sample depth groups (Appendix D.15F). The *hdrA* gene had an average abundance of 15.15 ± 1.71% between 1 – 6 cm and an average abundance of 40.08 ± 2.05% between 7 – 36 cm (Figure 4.24G). AOV showed a statistical difference in *hdrA* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and borderline significant differences between pairs 2 – 4 and 2 – 5 (Appendix D.15G). The *cooS* gene had an average abundance of 2.32 ± 0.25% between 1 – 6 cm and 6.04 ± 0.23% between 7 – 36 cm (Figure 4.24H). AOV showed a statistical difference in *hdrA* abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups, and between pairs 2 – 4 and 2 – 5 (Appendix D.15H).

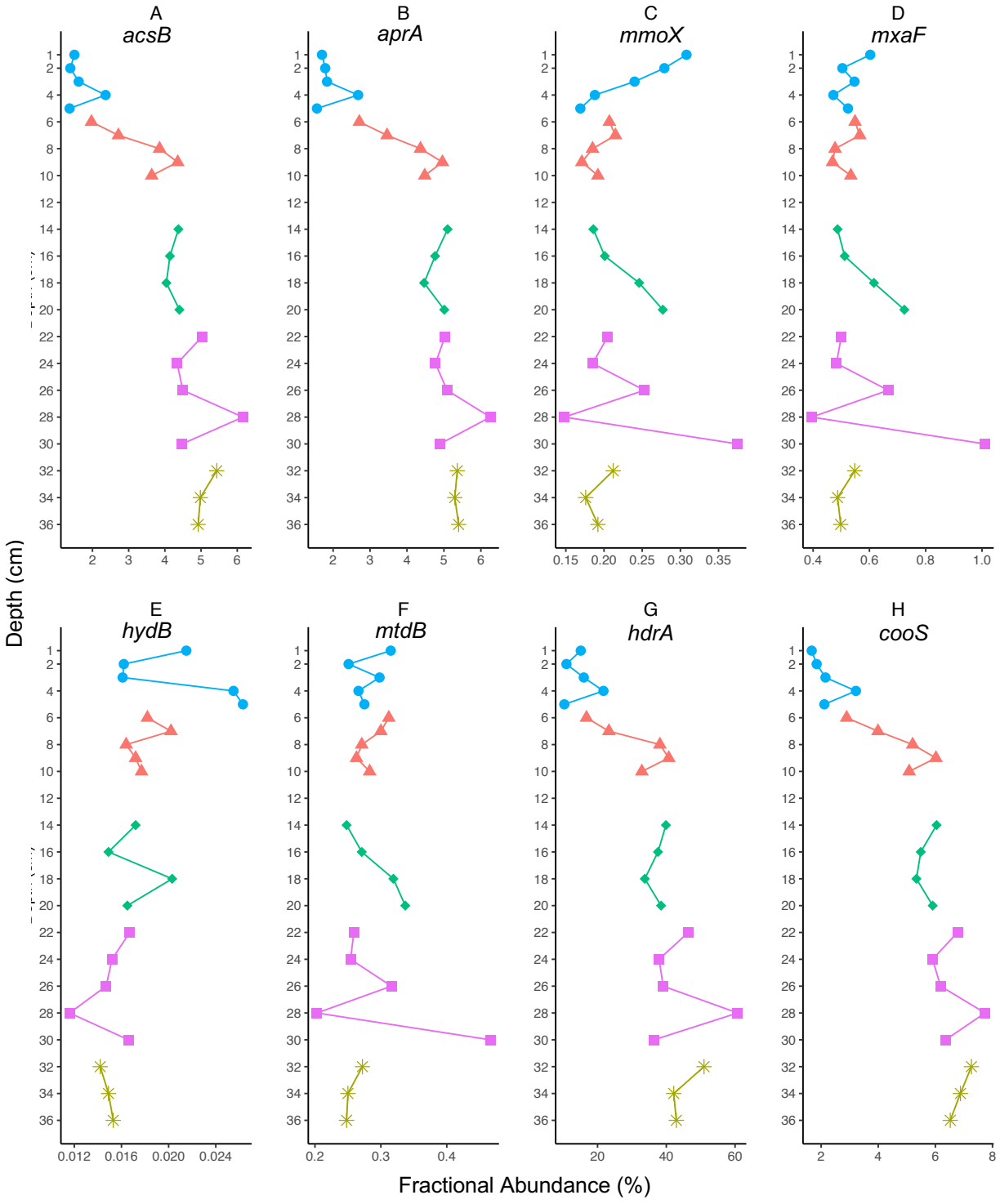


Figure 4.24. Mud Patch Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS*.

4.3.7 Geochemistry

4.3.7.1 *Ravenglass Depth Core Iron and Manganese Geochemistry*

Porewater Mn (Figure 4.25A) was most abundant at the top 1 – 5 cm of the core, ranging from 103 μM at 2 cm to 16 μM at 5 cm, and decreased to an average of 1.36 ± 0.3 μM between 6 – 28 cm. AOV showed a statistical difference in porewater Mn concentrations across sample groups, and Tukey HSD showed a statistical difference between sample group 1 and all other groups (Appendix D.16A). Porewater Fe (Figure 4.25B) was also most abundant in the upper 1 – 4 cm, increasing from 18.75 μM at 1 cm and peaking at 152.23 μM at 3 cm. At 5 cm, this decreased to 45.54 μM and at 6 cm to 18.3 μM. At 7 cm, it increased to 100 μM, and decreased from 8 – 28 cm to an average of 38.54 ± 6.2 μM. AOV showed a statistical difference in porewater Fe concentrations across all sample groups, with Tukey HSD showing a statistical difference between sample group pair 1 – 4 (Appendix D.16B). Solid Mn (Figure 4.25A) (Mn/Al ratio) was 0.013 at 1 cm, and between 2 – 14 cm had an average of 0.00969 ± 0.0005. At 16 cm, this increased to 0.0186, and between 18 – 28 cm, ranged from 0.0159 at 18 cm to 0.0096 at 24 cm (average 0.0129 ± 0.0011). AOV showed a statistical difference in Mn/Al values across sample groups with Tukey showing a statistical difference between sample group pair 1 – 3 (Appendix D.16D). Solid Fe (Figure 4.25B) (Fe/Al) ranged from 0.575 at 1 cm to 0.582 at 16 cm (average 0.569 ± 0.003). At 18 cm is 0.56 and decreased to an average 0.533 ± 0.004 between 20 – 28 cm. AOV showed a statistical difference in Fe/Al values across all sample groups, with Tukey showing statistical differences between sample group 4 and all other groups (Appendix D.16E).

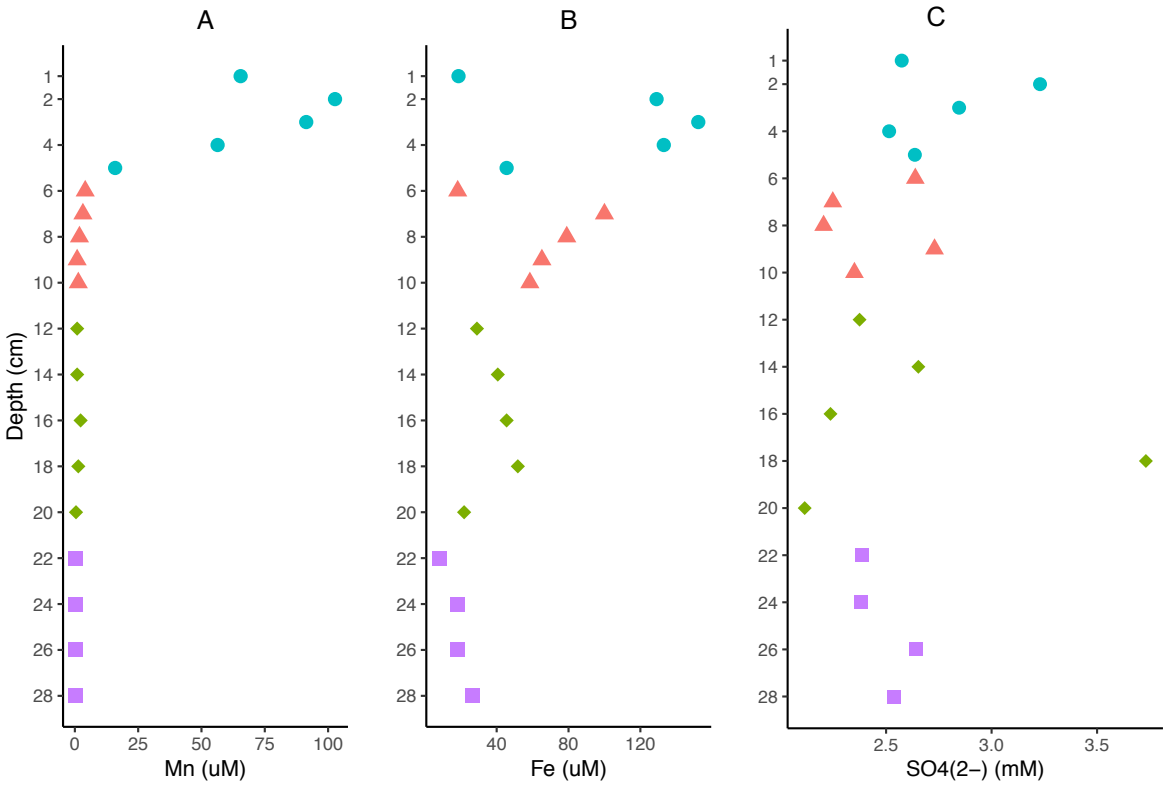


Figure 4.25. Ravenglass Depth Core: Depth profiles of porewater Mn (A)., Fe (B). concentrations in μM, and porewater sulfate (SO_4^{2-}) concentration (C). in mM from ion chromatography (IC).

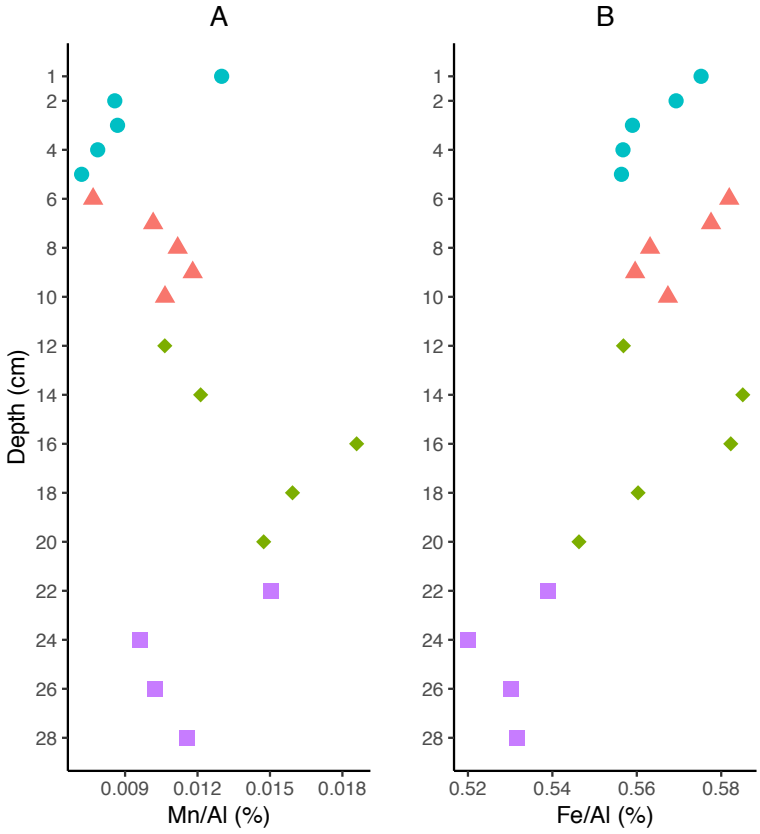


Figure 4.26. Ravenglass Depth Core: Depth profile of XRF Mn (A). and Fe (B). values normalised by aluminium (Al).

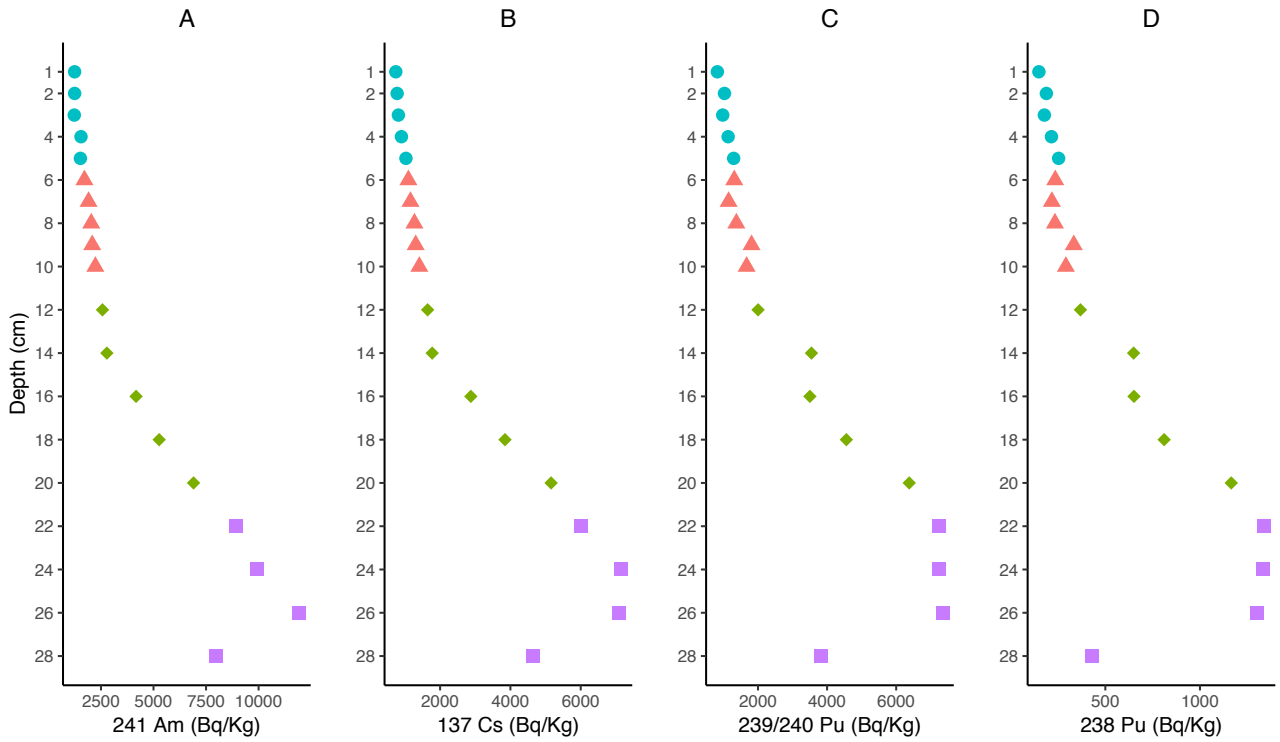


Figure 4.27. Ravenglass Depth Core: Depth profiles of activity of four radionuclide isotopes. X axis denotes activity in Bq/Kg and Y axis depth in cm. (A). ^{241}Am; (B). ^{137}Cs; (C). $^{239/240}$Pu; (D). ^{238}Pu.

4.3.7.2 *Ravenglass Depth Core Ion Chromatography*

Sulfate concentrations derived from ion chromatography (IC) did not have a depth-related pattern (Figure 4.25C), with an average concentration of 2.58 ± 0.09 mM across the entire core. AOV did not show a statistical difference in concentrations across the sample groups (Appendix D.16C).

4.3.7.3 *Ravenglass Depth Core Radionuclide Distribution*

^{241}Am,^{137}Cs, and $^{239/240}$Pu and ^{238}Pu activity distributions all increased with depth (Figure 4.27A, 28B, 28C, and 28D respectively). ^{241}Am increased between 1 – 14 cm from 1245.9 Bq/kg at 1 cm to 2781.5 Bq/kg at 14 cm (average = 1843.5 ± 149.7 Bq/kg). Between 16 – 28 cm, the average activity was 7874.6 ± 1016.4 Bq/kg, peaking at 11944.3 Bq/kg at 26 cm. ^{137}Cs had an average activity of 1159.5 ± 96.9 Bq/kg between 1 – 14 cm, and this increased to 5250.4 ± 609 Bq/kg between 16 – 28 cm, peaking at 7145.3 Bq/kg at 24 cm. $^{239/240}$Pu had an average activity of 1510.9 ± 210.9 Bq/kg between 1 – 14 cm, which increased to 5727.6 ± 645.2 Bq/kg between 16 – 28 cm. ^{238}Pu was lower than $^{239/240}$Pu, with an average activity of 276.2 ± 38.5 Bq/kg between 1 – 14 cm, and increased to 1004.9 ± 140.2 Bq/kg between 16 – 28 cm. AOV showed a statistical difference in all radionuclide activity values across all

sample groups, with Tukey HSD showing a statistical difference between sample group 4 and all other groups, and between pairs 1 – 3, and 2 – 3 (^{238}Pu activity only for the latter pair) (Appendix D.16F–I).

4.3.7.4 *Mud Patch Depth Core Iron and Manganese Geochemistry*

Porewater Mn (Figure 4.28A) is also most abundant in the upper 1 – 4 cm, with an average of 296.6 ± 21.5 μM, decreasing between 5 – 8 cm from 178.2 to 88.2 μM (average 150.8 ± 24.8 μM). Between 9 – 28 cm, the average decreases to 46.88 ± 4.88 μM. AOV showed a statistical difference in porewater Mn concentrations across all sample groups, with Tukey HSD showing a statistical difference between sample group 1 and all other groups, in addition to group pair 2 – 5 (Appendix D.17A). Porewater Fe (Figure 4.28B) generally decreases with depth, peaking at 2 cm with 226.3 μM, dropping to 3.13 and 5.8 μM at 4 and 5 cm. Between 6 – 28 cm, the average abundance is 103 ± 8 μM. AOV did not show a statistical difference in porewater Fe concentrations across sample groups (Appendix D.17B). Solid Mn (Figure 4.29A) is most abundant in the top 3 cm of the core, peaking at 0.0104 at 3 cm. This decreases to an average 0.0078 ± 0.00016, with a low of 0.0064 at 6 cm.

Solid Fe (Figure 4.29B) was most abundant from 1 – 12 cm with an average ratio of 0.463 ± 0.012, including a drop at 6 cm to 0.341, similar to as seen in the solid Mn. From 14 – 28 cm, the average decreases to 0.411 ± 0.0035. AOV showed a statistical difference for both Mn/Al and Fe/Al values across the sample depth groups, and Tukey HSD showed statistical differences between group 1 and all other groups for Mn/Al (Appendix D.17D) and between 1 – 4 for Fe/Al (Appendix D.17E).

4.3.7.5 *Mud Patch Depth Core Ion Chromatography*

Sulfate concentrations derived from ion chromatography (IC) (Figure 4.28C) appeared to increase between 0 – 10 cm, from 3.31 μM at 1 cm to 4.68 μM at 10 cm (average = 3.71 ± 0.19 μM), which then decreased to an average of 2.86 ± 0.05 μM between 12 –36 cm. AOV showed a statistical difference in SO_4^{2-} concentrations across all sample groups, with Tukey HSD showing a statistical difference between sample group 2 and all other groups (Appendix D.17C).

C

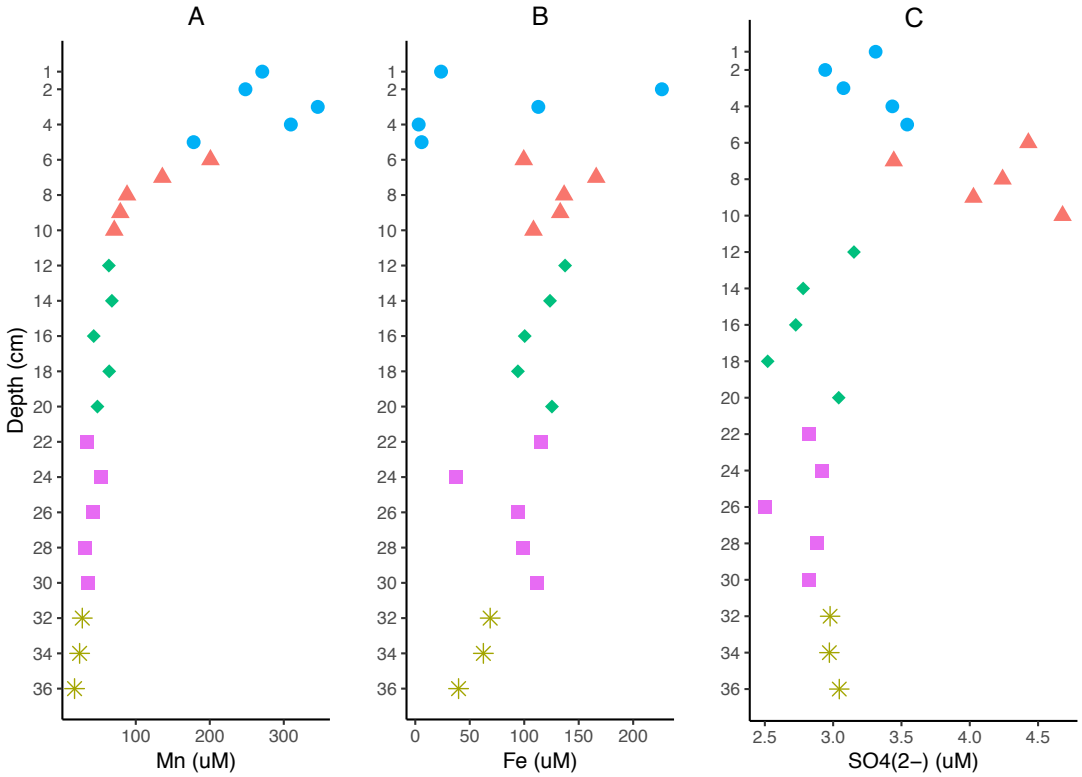


Figure 4.28. Mud Patch Depth Core: Depth profiles of porewater manganese (Mn) (A)., iron (Fe) (B). concentrations as a function of depth, and porewater SO_4^{2-} concentration (C). in μM as a function of depth from ion chromatography (IC).

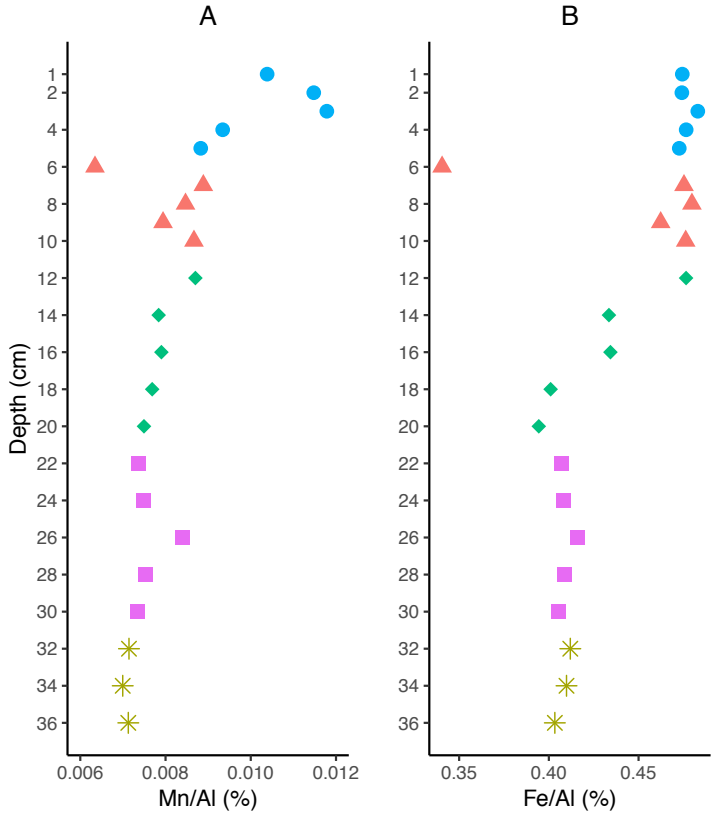


Figure 4.29. Mud Patch Depth Core: Depth profiles of XRF Mn (A). and Fe (B). values normalised by aluminium (Al).

4.3.7.6 *Mud Patch Depth Core Radionuclide Distribution*

^{241}Am, ^{137}Cs, and $^{239/240}$Pu and ^{238}Pu all showed broadly scattered distributions throughout the Mud Patch core, with some accumulation at depth (Figure 4.30A, 32B, 32C, and 32D respectively). ^{241}Am increased with depth from 205.1 Bq/kg at 1 cm to 382.3 Bq/kg at 34 cm, with an average activity of 338.2 ± 17.6 Bq/kg, and at 36 cm, the activity was 744.2 Bq/kg. ^{137}Cs abundance increased from 122.7 Bq/kg at 1 cm to a peak of 210.8 Bq/kg at 18 cm (average = 161.9 ± 7.6 Bq/kg) and remained steady between 20 – 36 cm with an average activity of 188.4 ± 3 Bq/kg. The $^{239/240}$Pu and ^{238}Pu fractions had similar distribution profiles with depth, broadly increasing between 1 – 24 cm, from 140.8 and 23.6 Bq/kg at 1 cm to 318 and 51.2 at 22 cm (average = 268.4 ± 15.3 and 45.8 ± 2.1 Bq/kg) respectively. At 26 cm, both Pu fractions increased in activity to 551.3 and 92.8 Bq/kg respectively and remained steady between 28 – 36 cm with average activities of 320.1 ± 6.1 and 55.6 ± 0.5 Bq/kg respectively. AOV showed a statistical difference in all radionuclide activity values across all sample groups (Appendix D.17F–I). Tukey HSD showed a statistical difference in: ^{241}Am values between group pairs 1 and 3, 4, and 5, and pair 2 – 5 (Appendix D.17F); ^{137}Cs between group 1 and groups 3, 4, and 5, and between group 2 and groups 3, 4, and 5 (Appendix D.17G); and both Pu fractions between group pair 1 – 4 (Appendix D.17H and 19I).

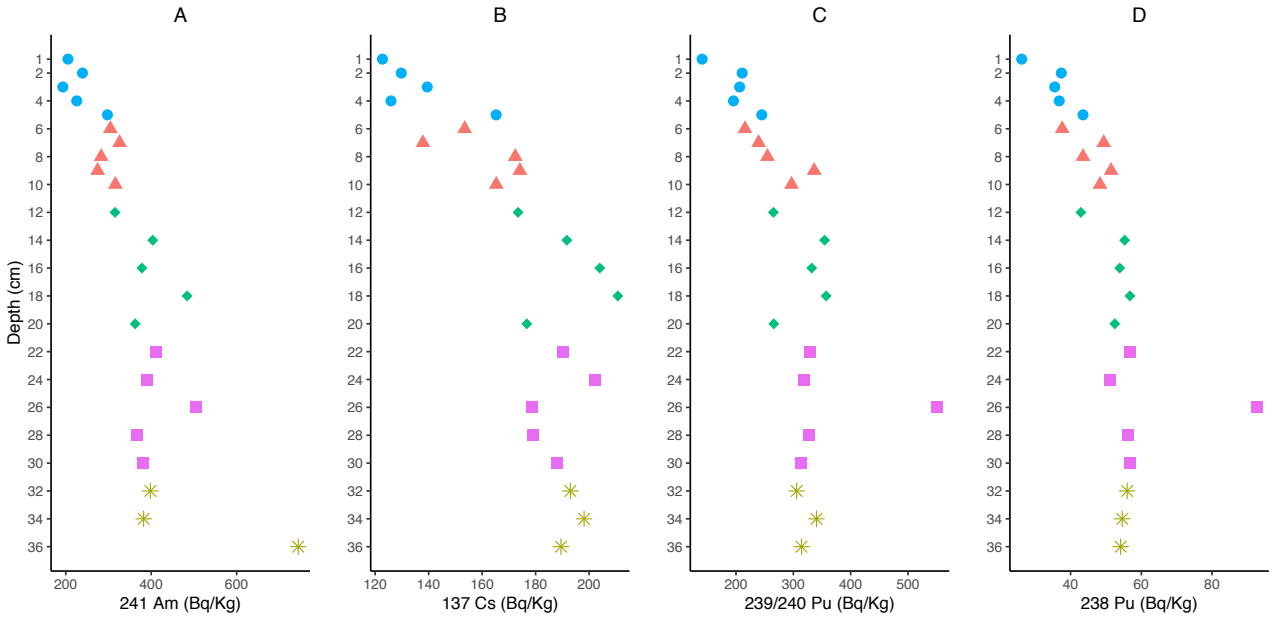


Figure 4.30. Mud Patch Depth Core: Depth profiles of activity of four radionuclide isotopes. X axis denotes activity in Bq/Kg and Y axis depth in cm. (A). ^{241}Am; (B). ^{137}Cs; (C). $^{239/240}$Pu; (D). ^{238}Pu.

4.3.8 Environmental Fitting of Geochemical Data

Correlation of environmental variables to the DCA plot of the Ravenglass microbial community assemblages was performed. TOC and dissolved Mn^{4+} appeared to correlate with the surface samples (1 cm and 2 – 7 cm respectively), while the radionuclides all correlated with the deeper samples (18 – 28 cm) and bulk Mn, and porewater and bulk Fe did not correlate with any samples (Figure 4.31).

For the Mud Patch depth core, only bulk Mn appeared to be correlated with the 1 – 3 cm and porewater Mn did not appear to correlate with any particular samples, while no other geochemical variable correlated with the DCA axes (Figure 4.32).

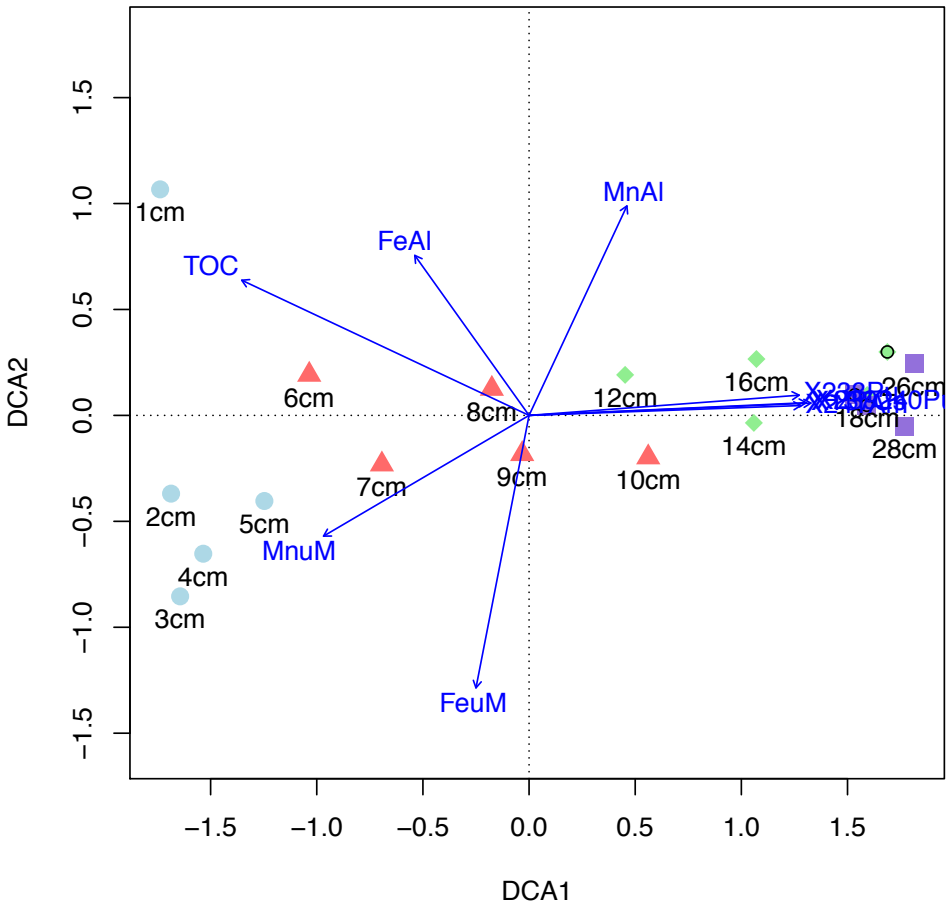


Figure 4.31. Ravenglass Depth Core: DCA plot with environmental variable data represented *via* environmental fitting correlation, represented by the blue arrows. Direction of arrows shows direction of increasing gradient and length of arrow is proportional to the correlation between the variable and the ordination axis.

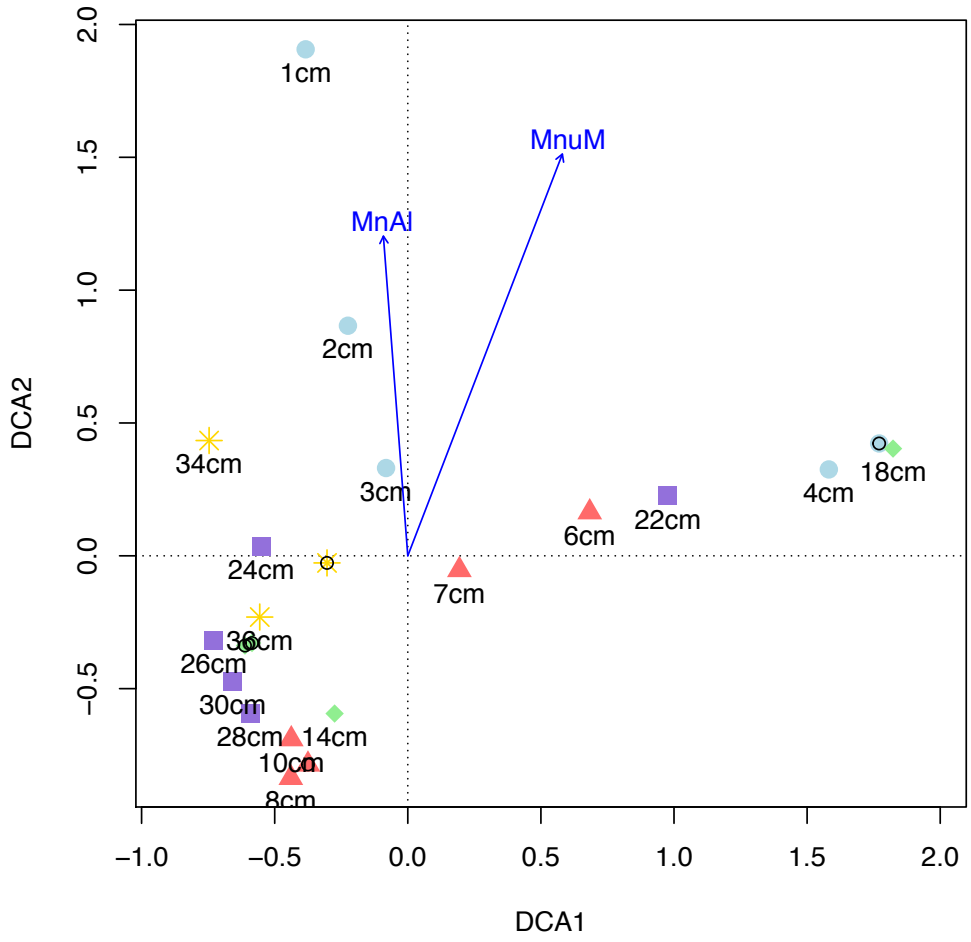


Figure 4.32. Mud Patch Depth Core: DCA plot with environmental fitting added.

4.4 Discussion

4.4.1 Contrasting the Ravenglass and Mud Patch Depth Cores

The two depth cores appeared to have very different microbial community assemblages as a function of depth, with neither core strictly adhering to the normally expected thermodynamically and diffusion controlled vertical stratification of redox processes in sediments (Thauer *et al.*, 1977). Both cores contained evidence of community assemblage being controlled by the surrounding environmental conditions. The Ravenglass core was taken from the mud flats close to the estuary, subject to daily tidal flows, direct inputs of carbon from the surrounding terrestrial environment, and periodic exposure of surface sediments to the air. The Ravenglass core appeared to have two main microbial communities associated with the surface and the lower depth samples, possibly due to the complex hydrological nature imposed by tidal flows. The Mud Patch represented a fully marine environment where marine primary production, subsurface currents and diffusional processes contribute to inputs. The Mud Patch had some evidence for vertical stratification of microbial community and function but appeared to be severely limited in succession and defined more by a heterogeneous mix of aerobic and anaerobic processes, as well as there being evidence of bioturbation by indigenous macrofaunal mixing the microbial communities.

One potential driver of the differences between these two sites is the substantial differences in the overall levels of organic carbon available to drive microbial processes and possibly play a role in retaining metals. Ravenglass which has between 4 – 7% TOC (Figure 4.1B) and, based on quantification of DNA extractions (Figure 4.1A), substantially greater levels of biological activity, is characterised by a clear pattern of depth-related succession in geochemistry, microbial community composition and function and can be considered a more ‘classical’ diagenetic environmental depth core. The Mud Patch with relatively low levels of total carbon (0-1.2%) (Figure 4.3B) likely reflects much lower inputs of detrital organic carbon and hence lower levels of biological activity. Possibly as a result of this, the Mud Patch is generally a heterogeneous mix of geochemical conditions and microbial taxa, with only minor evidence of redox succession after ~ 7 cm, albeit critically with no direct evidence of the classical depth-related redox succession. This lack of redox succession in the Mud Patch is certainly consistent with low levels and likely refractory nature of the carbon in this system which will

restrict the supply of electron donor to drive biotic (O_2, NO_x^-, Fe^{3+}, and SO_4^{2-}) abiotic organic carbon coupled redox processes.

Alpha diversity indices showed clear and consistent differences in microbial community structures between the two cores. All three indices of diversity measured were consistently higher in the Ravenglass depth core (Figure 4.7), indicating a more species rich, phylogenetically diverse, and even microbial community at Ravenglass than the Mud Patch. Furthermore, all three indices followed the same pattern with depth in the Ravenglass depth core of an increase between 1 – 5 cm group and 6 – 10 cm group, followed by a decline to 28 cm, evidence of progressive community structural changes with depth within the core (Appendix D.5). The increased richness, diversity, and evenness exhibited within the Ravenglass core versus the Mud Patch core may simply be another reflection of the differing carbon availability levels at the sites, implying that the higher levels of TOC at Ravenglass enable the site to host a more diverse microbial community due to higher respiration by the microbial community (Cleveland *et al.*, 2007). Interestingly, both the Ravenglass and Mud Patch depth cores had lower alpha diversity indices values when compared to the Needle's Eye (Chapter 3) and Loch Etive (Chapter 5) depth cores. This would be expected at Needle's Eye due to the terrestrial nature and much higher levels of TOC at the site, but the Loch Etive cores are also marine sediment cores. Phylogenetic diversity is especially low in Ravenglass and Mud Patch compared to Needle's Eye and Loch Etive, indicating that the communities at Ravenglass and Mud Patch are less evolutionarily diverse and likely less functionally diverse. Due to the well-documented OTU inflation issue in many older OTU-picking methods massively overestimating alpha diversity values (Edgar, 2017), it is not viable to compare the values presented for Ravenglass and Mud Patch with similar studies due to the more stringent methods employed here.

Beta diversity of microbial community assemblages did present an interesting insight into the community assemblages between the two cores. Ordination using PCoA, nMDS, and DCA of the ASV microbial communities – that is, comparing the ASV representative sequence tables between the two cores as derived from DADA2 – appeared to show complete dissimilarity in ASV assemblages. However, when taxonomic classifications were compared at various classification levels, there was still community dissimilarity in taxonomic assemblages between the two sites resulting in complete separation of samples, but there was a depth-related

distribution of samples of both cores (Figure 4.8). Furthermore, the Ravenglass depth core had clear, consistent depth-related distribution of microbial community assemblages (Figures 4.9 and 4.10) whereas the Mud Patch was defined by community homogeneity with the exception of several individual depth increments (Figures 4.11 and 4.12).

However, a site-specific assessment of how the specific conditions, biology, and geochemistry spatially co-vary at the sites enables an assessment of the individual processes occurring. This assessment is considered in the context of the physical setting (coastal estuary or deep water fully marine) and allows a prediction about the long-term transport and fate of radionuclide contaminants in these environments.

4.4.2 The Ravenglass Depth Core

Though the beta diversity ordination plots show a progressive change in microbial community assemblage as a function of depth in the Ravenglass core, taxonomic and functional assemblage of the core appeared to show the core actually divided into two main halves with a transition zone in between, defined here as the 1 – 4 cm, 5 – 9 cm, and 10 – 28 cm regions. These regions are likely constructs of the complex tidal hydrology of the site, where the tide rises in the estuary bidaily and likely causes significant mixing of the site in an upwards trajectory. Based on the microbial taxonomic and functional community assemblages, it appears that the tidal mixing effects the 10 – 28 cm region, whereas 1 – 4 cm appeared to have very spatially-short redox stratification, and 5 – 9 cm section represented the cross between the two. Overall, Ravenglass radionuclide profiles reflected historical releases from Sellafield and have been deposited at depth, and in the case of plutonium (Pu), is likely in the same Pu^{4+} state at depth as it was when released.

4.4.2.1 *The 1 – 4 cm Region of the Ravenglass Depth Core: Rapid Succession*

It is first worth noting that while TOC was highest in this region than any other of either site, DNA concentrations actually increased from 1 – 4 cm (Figure 4.1). Alpha diversity was also consistently lower in this region than other regions, implying a less species rich, phylogenetically diverse, and even microbial community compared with the 5 – 10 cm groups (Figure 4.7, Appendix D.5). Lastly, Beta

diversity ordination showed that the community assemblage change between 1 – 4 cm was linear with regards to the principal component (PC), nMDS, and DCA axes, but not in the same direction as the remaining samples, indicating that the community assemblages changed in this region in a different manner to that observed with the rest of the core (Figures 4.9 and 4.10).

'End point' PCRs appeared to show the denitrifying *nirK* and *dsrA* genes at their most abundant in this region of the core, while the bacterial and archaeal *amoA* genes were lower in this region than following regions (Figure 4.5, Appendix D.3). This is in contradiction to the normally expected stratification of redox profiles, where oxygen diffuses from the surface, and appears to imply the presence of relatively reducing conditions from the surface.

Taxonomically, this region was actually comprised of a range of coexisting aerobic and anaerobic taxonomic groups that had a progressive and incremental change in abundance with depth (Figure 4.13). For instance, *Woeseia* spp. was the most abundant genus and is of interest as this group is considered aerobic and chemoheterotrophic albeit that all genomes analysed also contain a truncated denitrification pathway (*nirS* and *norB*) as well as evidence of putative sulfur and hydrogen oxidation (Du *et al.*, 2016; Mußmann *et al.*, 2017). This group dropped in abundance slightly between 1 – 4 cm. Likewise, the strictly aerobic chemoheterotrophic *Sandaracinaceae* (Mohr *et al.*, 2012) decreased between 1 – 4 cm (Figure 4.14C) as did the *Saprospiraceae*, a family associated with aerobic chemoheterotrophy (McIlroy and Nielsen, 2014) (Figure 4.14F). Conversely, the facultatively aerobic chemoheterotrophic iron and nitrate-reducing

Ardenticatenales family was also present in this region (Kawaichi *et al.*, 2013) in the form of two distinct ASVs that increased between 1 – 4 cm and the strictly anaerobic fermenting *Anaerolineaceae* (Yamada *et al.*, 2006) were abundant in this region and indeed throughout the core. In addition, two strictly anaerobic sulfate-reducing *Desulfobulbaceae* (Kuever, 2014) both peaked at 3 cm (Figures 4.14A and 4.14E). The facultatively anaerobic, nitrogen-fixing *Prolixibacteraceae* (Huang *et al.*, 2014) increased from 1 – 4 cm (Figure 4.14G).

Tax4Fun profiles of inferred gene abundances concurred with the 'end point' PCR abundances for *amoA* and *nirS* (low abundance) and *dsrA* (high abundance), while *nirK* profiles contradicted one another (Figure 4.17). The *dsrA* abundance increased progressively from 1 – 4 cm. Nitrogen cycling genes showed mixed profiles

of abundances in this region. While the *narG* increased progressively between 1 – 4 cm (and continued to increase until 7 cm), the *napA*, *nrfA*, and *nifD* all peaked in this region (Figure 4.19). The *norB* and *nosZ* appeared to decrease in this region. Of the other functional genes targeted in the Tax4Fun inferred profiles, anaerobic genes including the *acsB*, *aprA*, *hdrA*, and *cooS* all increased between 1 – 4 cm (Figure 4.20).

These profiles indicate a region that is characterised by a rapid transition from oxic to putatively sulfate-reducing conditions, likely driven by undisturbed environmental conditions (*i.e.*, away from the tidal zone) and relatively sufficient labile carbon levels (on the basis that all abundant species are chemoheterotrophic). The co-location of strictly aerobic and anaerobic taxa indicates that this spatially-short stratification is likely to be in the form of zones of anaerobic conditions within an otherwise aerobic environment driven by the diffusion of oxygen from the surface. However, a lack of oxidising genes such as the *amoA*, or *mmoX* implies that oxygen is limited in this region, although this may also be due to a preference for chemoheterotrophic aerobic respiration. The observed increase in sulfate-reducing taxa and genes between 1 – 4 cm, and the presence of the iron-reducing *Ardenticatenales*, implies that there is likely microbial reduction of Fe^{3+} and SO_4^{2-}. The increase of anaerobic carbon fixating genes (*acsB* and *cooS*) is further evidence of anaerobic conditions. The increase in DNRA and nitrogen-fixation genes is also potential evidence for anaerobic conditions (Gray and Head, 2005; Leschine *et al.*, 1988) and was associated with anaerobic conditions in the Needle's Eye depth core (Chapter 3). The lack of *mcrA* or *Euryarchaeota* methanogenic archaea in this region (or in the entire core) implies a lack of methanogenesis.

Geochemical analysis appeared to strongly suggest the reduction of solid phase Mn^{4+} and Fe^{3+} into aqueous Mn/Fe^{2+}, and the reduction of porewater sulfate, in this zone. Solid phase Mn and Fe decreased between 1 – 3 cm (Figure 4.26), while porewater Mn and Fe increased in this zone (Figure 4.25), indicating the reduction of Mn and Fe, probably *via* microbial bioreduction. Porewater sulfate concentrations peaked at 2 cm and decreased between 2 – 4 cm, implying reduction of sulfate (Figure 4.25C). This corroborates with a previous measurement of E_h at the site that showed a sub-oxic reduction potential of 180 – 230 mV (Marsden *et al.*, 2006).

4.4.2.2 *The 5 – 9 cm of the Ravenglass Depth Core: Transition Zone*

This cross-over section between the upper more classical 1 – 4 cm (described above) and the unexpectedly complex lower 10 – 28 cm samples is characterised by a decrease in the range of anaerobic taxa and genes, and the emergence of another mixed aerobic and anaerobic communities. This region had the highest alpha diversity index values of the two cores (but not significantly so), while beta diversity showed these samples to be relatively spread out on the ordination plots compared to the other two described regions, indicating a greater amount of community dissimilarity between samples. TOC and DNA values continued to decline in this region.

Taxonomic and functional gene profiles had complicated distribution patterns and varied on the cm scale. There were no clearly-defined depth related decrease or increase in abundances in this region. 'End point' PCR showed the archaeal and bacterial *amoA*, and *nirS* all increased though not significantly, and the *nirK* and *dsrA* did decrease, although only the *dsrA* decreased significantly (Figure 4.5, Appendix D.3).

Many of the taxa described for the near surface region were also found at this lower depth. For instance, the aerobic *Woeseia* spp., and anaerobic *Ardenticatenales*, and *Anaerolineaceae* were present but did not significantly change in their abundances (Figure 4.13, Appendix D.8). The aerobic *Sandaracinaceae* did not statistically change in abundance compared to 1 – 4 cm but did decrease between 5 – 9 cm (Figure 4.14C) and the aerobic *Saprospiraceae* did significantly decrease (Figure 4.14F and Appendix D.8P). More obviously, the two *Desulfobulbaceae* ASVs which predominated at shallower depths statistically decreased in this region. Of particular note, was the emergence of the ammonia oxidising *Thaumarchaeota* genus *Candidatus Nitrosopumilus* (Walker *et al.*, 2010) in this region (Figure 4.14I), as was the putative methane-cycling (either methanotrophic or methanogenic) *Bathyarchaeia* (Evans *et al.*, 2015) (Figure 4.14J). This decrease in chemoheterotrophic taxa (*Sandaracinaceae*, *Saprospiraceae*) and the emergence of chemolithotrophic archaea is the first tentative indication of changing carbon levels impacting the microbial community assemblage.

Tax4Fun showed that *amoA* genes statistically increase in this region similar to the 'end point' PCR, and this coincided and was corroborated by the emergence of the *Nitrosopumilaceae* (Figure 4.17A, Appendix D.10A). The *dsrA* statistically

decreased in this region, as it did in the 'end point' PCR, coinciding with the decrease in *Desulfobulbaceae* (Figure 4.17D and Appendix D.10D). Nitrogen cycling genes continued to show mixed profiles of abundances with depth. The *narG* increased from 5 – 7 cm but decreased sharply from 7 – 8 cm (Figure 4.19A). The *napA* decreased from 5 – 7 cm but increased from 7 – 8 cm (Figure 4.19B). The *norB* and *nosZ* genes increased from 5 – 9 cm (Figures 4.19C and 4.19D). The *nrfA* and *nifD* decreased from 5 – 9 cm (Figures 4.19E and 4.19F). Other target genes showed a more consistent pattern. The anaerobic *acsB*, *aprA*, *hdrA*, and *cooS* all statistically decreased (Figure 4.20, Appendix D.12), while the oxidising *mmoX,* the other methane-oxidising *mxaF*, and the methanotrophic *mtdB* increased in this region. The source of this methane is not obvious due to the lack of methanogenic processes in this core. The NiFe hydrogenase, associated with a range of H_2 metabolic processes including methane cycling, increased in this region (Figure 4.20E).

The microbial taxonomic and functional profiles of this region imply mixed and disturbed conditions, but conditions that are overall less putatively strictly anaerobic as observed at 3 – 4 cm, and potentially more oxidising as shown by the emergence of *Nitrosopumilaceae* and *amoA*, and *mmoX* taxa and genes.

Geochemical analysis appeared to corroborate this picture of mixed aerobic and anaerobic heterogeneity. Solid phase Mn^{4+} increased slightly in this region, while soluble porewater Mn^{2+} concentrations were negligible (and for the rest of the core), potentially indicating that any reduced Mn^{2+} is being removed from solution and is precipitating into the solid phase, either *via* oxidation or sorption. Solid phase Fe, presumably oxidised Fe^{3+}, increased slightly between 5 – 7 cm, while porewater Fe, presumably reduced Fe^{2+}, decreased between 5 – 9 cm. The removal of reduced Fe is potentially *via* precipitation of Fe (oxyhydroxides) or sorption to sediment surfaces. Porewater sulfate levels had no discernible pattern in this region, indicating no further reduction of sulfate. The conditions in this region are probably not fully oxidising but are certainly less reducing than the 0 – 5 cm region.

4.4.2.3 *The 10 – 28 cm Region of the Ravenglass Depth Core: Perturbed System*

This region of the core is characterised by a heterogeneous mix of oxidising and aerobic, and reducing and anaerobic microbial taxa and functional genes, with the community becoming more homogenous with depth. In this region, DNA and TOC concentrations were at their lowest and did not decrease with depth within this

region. Alpha diversity index values decreased and saw statistical decreases from the 6 – 10 cm group to the 22 – 28 cm group (Figure 4.7 and Appendix D.5). Beta diversity showed that the microbial community assemblages change more between 10 – 16 cm, and that the 18 – 28 cm region was much more homogenous in community assemblage (Appendix D.6).

'End point' PCR results showed low abundance and no pattern with the archaeal *amoA*, but the bacterial *amoA* and *nirK* did decrease significantly with depth. The *nirS* and *dsrA* decreased but not significantly.

This region was characterised by the decrease in several of the dominant taxa identified in the upper regions and emergence of different taxa. For instance, the *Woeseia*, *Desulfobulbaceae*, and *Saprospiraceae* all decreased to significantly lower levels. Many of the most abundant taxa that emerged in this region are poorly characterised, including the δ-*proteobacteria* MBNT-15 group, not yet cultured and with no known functions, the *Gemmatimonadetes*, and the *Pirellulaceae*. Interestingly, the *Anaerolineaceae* and *Ardenticatenales* persisted throughout. Taxa with known functions did emerge in this region, including the nitrogen cycling (DNRA and denitrification) and sulfate-reducing *Dadabacteriales*, which are also predicted to be involved with ammonia oxidation with *Thaumarchaeota* by supplying ammonia from DNRA (Hug *et al.*, 2016), the methanol oxidising *Methyloceanibacter* spp. (Takeuchi *et al.*, 2014), and the archaeal *Nitrosopumilaceae* and *Bathyarchaeia* increased.

Tax4Fun showed all the nitrogen cycling genes decrease in abundance, as well as the other anaerobic target genes, while the *mmoX*, *mxaF*, and *mtdB* all increased.

Based on the microbial community and functional assemblages, it is not possible to define with certainty the prevailing biogeochemical cycling and redox conditions of the core. The emergence of methane-oxidising genes and taxa implies the presence of both methane and oxygen, and the presence of ammonia oxidising genes and taxa implies the presence of ammonia and oxygen. However, it is possible to describe what the core does not have, which is the abundance of nitrogen cycling and sulfate reduction observed at the upper regions of the cores, although, the presence of the *Ardenticatenales* throughout implies the reduction of Fe may be occurring (Kawaichi *et al.*, 2013).

Geochemical analysis generally implied further evidence of sub-oxic conditions. Solid phase Mn and Fe did decrease implying the removal of oxidised Fe and Mn probably *via* microbially-mediated bioreduction, while the absence of increasing porewater Mn/Fe expected from the bioreduction process is likely due to the precipitation of Fe^{2+} as Fe (oxyhydroxides) or sorption to sediment surfaces. Porewater sulfate levels did not change in this region, indicating a lack of sulfate reduction, or potentially continual replenishment by oxidation of H_2S where sulfate is reduced in anaerobic microniches (Jørgensen 1977). The radionuclide levels all increased in this region, with their maximums found at ~ 22 – 26 cm (Figure 4.27), likely reflecting the historical release of these radionuclides from Sellafield in the 1970s having been deposited and sedimented, at a rate of around 6.8 mm yr^{-1} (Ray *et al*., in prep.).

4.4.3 The Mud Patch Depth Core

The microbial community taxonomic and functional profiles derived from the Mud Patch are very different from that of the Ravenglass depth core, and as such, the inferred geochemical conditions and predicted fate of radionuclides in the system is entirely different. The Mud Patch is a fully marine system and represents the outer extent of the Irish Sea tide. It has been documented that the Mud Patch acts as a transient sink for Sellafield-derived radionuclides, where instead of radionuclides being deposited and sedimented at depth, they are reworked by macrofauna and redistributed into the water column (Kershaw *et al.*, 1983) where they are then carried by the tide back to the coastlines of the UK and Ireland (McCartney *et al.*, 1994; Keogh *et al.*, 2007). Sedimentation rates at the Mud Patch are thought to range from 0.02 – 0.08 cm yr^{-1} (Ray *et al.* in prep.). All microbial and geochemical methods used reflected this dynamic and mixed nature of the Mud Patch core, with only minor evidence for the normally expected vertical stratification of redox processes after 7 cm observed.

The Mud Patch core shows little evidence of having accumulated or retained large amounts of the historical radionuclide releases from Sellafield. This lack of retention is perhaps not surprising given, as has already mentioned above, the obviously more oxidising character of these sediments largely controlled by low levels of organic matter. However, a depth related analysis of this core reveals some interesting characteristics relevant to radionuclide mobilisation and retention. For

instance, the Mud Patch depth core is characterised by specific ASVs dominating at particular depth samples; namely 1, 4, 5, 18, and 22 cm. One potential explanation for this sporadic variation in community composition against a back drop of normal, albeit mild, redox succession (see below) is the occasional physical and biological perturbation of these sediments. Such perturbations may be a result of macrofaunal activity such as burrowing marine worms (Kershaw *et al.*, 1983), however, there is no direct strong evidence to support this conclusion other than to note that these dominant species are all strictly aerobic species even at depth where it would normally be expected for oxygen to be depleted.

DNA and TOC concentrations were consistently low in the core, with TOC never exceeding 1.5% (Figure 4.3), no correlation between DNA and TOC values (Figure 4.4), and no statistical differences in DNA and TOC values as a function of depth (Appendix D.2). All alpha diversity index values were lower in the Mud Patch than any other core studied in this thesis (Figure 4.7) and did not change with depth as observed in other cores (Appendix D.5). Beta diversity ordination consistently showed several individual depth increments having a different microbial community assemblage, with 1, 2, and 3 cm samples forming one group, and 10 – 36 cm forming the other, with the exception of 4, 5, 7, 18, 22, and 30 cm all being different from the 10 – 36 cm cluster (Figures 4.11 and 4.12). Together, these results indicate a likely low-activity, low diverse microbial community that has several depth increments that are dominated by a small number of ASVs (hence the low evenness).

Regardless of sporadic variations in community composition there was, however, some evidence of depth-related stratification of microbial taxa and function, in the remaining depth samples broadly divided into 1 – 7 cm and 8 – 36 cm based on the microbial community similarity.

4.4.3.1 *The 1 – 7 cm Region of the Mud Patch Depth Core: Community Dominance*

The microbial community taxonomic and functional profiles reveal a very uneven community with several depth increments dominated by a few particular ASVs. ‘End point’ PCRs showed the region to be almost-exclusively comprised of the bacterial *amoA* genes (Figure 4.6B) but also evidence of the *nirS* increasing significantly in this region between 1 – 7 cm (Figure 4.6D, Appendix D.4D). The oxidase-positive, psychrophilic and halotolerant *Psychrobacter* spp. accounted for

40.07% of the 1 cm community assemblage abundance (Figure 4.15). An uncultured *Planococcaceae* ASV, an aerobic or facultatively anaerobic family (Shivaji *et al.*, 2014) 54.48%, 53.88%, and 16.33% at 4, 5, and 6 cm respectively. The mostly aerobic *Flavobacteriaceae* (McBride, 2014) ranged between 11.78% and 18.54% at 1 and 2 cm respectively. The strictly aerobic and chemoheterotrophic *Lutimonas* spp. (Yang *et al.*, 2007; Kim *et al.*, 2014) was confined to the 1 – 3 cm region (Figure 4.16A). The aerobic and chemoheterotrophic *Woeseia* spp. (Mußmann *et al.*, 2017) was also present in this region, and throughout the whole core (Figure 4.16B).

Tax4Fun inferred gene abundance profiles of the genes targeted by 'end point' PCR did not correlate well between the two methods (with the potential exception of the *dsrA* gene) (Figure 4.22). Of the 'end point' PCR target genes, the *amoA* had a mixed abundance profile in this region, generally increasing between 1 – 7cm (Figure 4.21A), the *nirK* decreased overall (Figure 4.21B), the *nirS* was lowest in this region than other regions of the core (Figure 4.21C), the *dsrA* increased progressively (Figure 4.21D). Similar to the 'end point' PCR, essentially no *mcrA* was detected (Figure 4.21E). This is likely due to the low abundances of archaea in general in this core, also reflected in the archaeal *amoA* 'end point' PCR band intensities.

Nitrogen cycling genes were all generally lowest in this region and increased between 1 – 7 cm, including the denitrifying *narG*, *napA*, *norB*, DNRA *nrfA*, and nitrogen-fixing *nifD*, while the final denitrifying *nosZ* decreased in this region (Figure 4.23). Of the other target genes, the anaerobic *acsB*, *aprA*, *hdrA*, and *cooS* were all lowest in this region, while the *mmoX* was generally low in abundance in general in this core but decreased between 1 – 5 cm (Figure 4.24).

These profiles indicate a region of the depth core that is almost entirely comprised of aerobic chemoheterotrophic processes, that are mixed, disturbed, and largely dominated by few individual taxa, but with evidence of gradual progressive community and functional change between 3 – 7 cm. Solid phase Mn decreased between 0 – 4 cm (Figure 4.31A) while porewater Mn increased in this zone (Figure 4.28A), likely as a result of microbial bioreduction (Froelich, 1979). Porewater Mn reduced after 4 cm, while bulk Mn changed little with depth, indicating the removal of reduced Mn^{2+} *via* sorption or precipitation as rhodocrosite. Solid phase Fe did not change in this zone, indicating a lack of microbial bioreduction, while porewater Fe concentrations were more heterogenous in distribution (Figure 4.28B, Appendix D.17B), possibly as a result of input of fresh ferric oxy(hydr)oxides to the surface of

the Mud Patch. Porewater sulfate actually increased in this region (Figure 4.28C, Appendix D.17C), indicating the introduction of sulfate likely from HS^- oxidation (Jørgensen, 1990). Radionuclide abundances are all consistently lower in this region than deeper at the core, potentially reflecting sedimentation of historical releases at depth in the core or mixing in this region suppressing accumulation. Due to the aerobic conditions in this upper region, it is unlikely that the indigenous microbial species are using the Sellafield-derived Pu as a TEA for anaerobic respiration.

4.4.3.2 *The 8 – 36 cm Region of the Mud Patch Core: Mixed Stratification*

Of the 'end point' PCR target genes, only the bacterial *amoA* persisted with depth. ASVs with increased abundance in this region include the sulfate-reducing family *Desulfobulbaceae* (Kuever *et al.*, 2014), the strictly anaerobic family *Anaerolineaceae* (Yamada *et al.*, 2006), and the strictly aerobic *Sandaracinaceae* (Mohr *et al.*, 2012), the *Thermoanaerobaculaceae* subgroup 10, the likely methanol oxidising *Methyloceanibacter* spp. (Takeuchi *et al.*, 2014), the sulfate-reducing δ-*proteobacteria* SEEP-SRB1, linked with anaerobic methane oxidation (Schrieber *et al.*, 2010; Timmers *et al.*, 2016; Skennerton *et al.*, 2017), as well as the *mcrA*-containing methane cycling *Bathyarchaeia* (Evans *et al.*, 2015). Similar to the Ravenglass depth core, it is not immediately clear why the methane-oxidising taxa are present. At 18 and 22 cm, the same uncultured *Planococcaceae* ASV that dominated at 3 and 4 cm accounted for 68.28% and 34.45% respectively. The *Psychrobacter* also re-emerged almost exclusively at 22, 24, 32, and 34 cm. Interestingly, the *Jeotgalibacillus* spp. and *Bacillus* spp. accounted for 5.74% and 4.49% at 22 cm exclusively. The *Jeotgalibacillus* is a genus in the *Planococcaceae* family and are considered aerobic (Srinivas *et al.*, 2016).

Tax4Fun appeared to show an increase in *dsrA* abundance, likely linked to the emergence of the *Desulfobulbaceae* and possibly δ-*proteobacteria* SEEP-SRB1 (Figure 4.21D) as well as the nitrite-reducing *nirS* (Figure 4.21C). The denitrifying *narG*, *napA*, and *norB* genes all increased and remained steady in this region, while the *nosZ* declined, indicating an incomplete denitrification pathway, while the *nrfA* and *nifD* increased, potentially indicating anaerobic conditions (Figure 4.23). The anaerobic genes, *acsB*, *aprA*, *hdrA*, and *cooS* were all most abundant in this region, while the methane-oxidising genes *mmoX*, *mxaF*, and *mtdB* were all low in abundance (Figure 4.24).

Geochemical analysis showed solid phase Mn decline slightly but not significantly alongside near-complete removal of porewater Mn, indicating the removal of reduced Mn^{2+} *via* sorption or precipitation as rhodocrosite. Solid phase Fe reduced between 12 – 18 cm and remained constant thereafter, indicating some bioreduction of Fe in this zone, while porewater Fe also generally decreased throughout this zone, albeit not statistically, possibly reflecting the removal of Fe^{2+} *via* precipitation of Fe (oxyhydroxides) or sorption to sediment surfaces. Porewater sulfate levels decreased after 10 cm, coinciding with the increase in *Desulfobulbaceae* and *dsrA* abundances (Figure 4.28C), indicating the reduction of sulfate after 10 cm. Interestingly, solid phase Mn and Fe decreased with depth, indicating the removal of oxidised Mn^{4+} and Fe^{3+}.

The increase in denitrification genes, and sulfate-reducing taxa and genes, as well as the reduction of porewater Mn, Fe, and sulfate levels, indicate that this region of the core must have the capacity for microbially-mediated Fe and sulfate reduction. The depth increments that observed large enrichments of *Planococcaceae* and *Psychrobacter* spp. are likely evidence of subsurface mixing caused by burrowing polychaetes introducing aerobic surface taxonomic groups (Kershaw *et al.*, 1983). The mixed aerobic and anaerobic microbial profiles is likely a reflection of the limited carbon availability resulting in low microbial activity. It is not by any means certain that the presence of certain microbial functional groups in the 16S rRNA gene sequencing means that function is definitely occurring. At depth in the Mud Patch depth core, there are likely microenvironments of some anaerobic function, but it is probable that any catabolic processes are carbon limited and slow – hence the lack of clearly defined stratification observed in other cores studied as part of this thesis.

4.4.4 The Fate of Radionuclides at Ravenglass and Irish Sea Mud Patch

Taxonomic and functional profiles constructed for the Ravenglass and Mud Patch depth cores produced mixed and inconclusive redox inferences, with little evidence for redox stratification as a function of increasing depth normally expected (Thauer *et al*., 1977). Specifically, that at Ravenglass, the 0 – 4 cm region was potentially sulfate reducing, 5 – 9 cm a transition zone of increased oxidation conditions, and the remaining 10 – 28 cm a heterogenous and perturbed mix of oxidizing and sub-oxic conditions, while the Mud Patch core appeared to be perturbed and mixed throughout with little evidence for redox stratification. In the

Ravenglass depth core, the radionuclides measured were consistently most abundant in the lower ~ 20 – 28 cm region of the core, in what the microbial profiles depicted to be of mixed aerobic and sub-oxic conditions. In the Mud Patch depth core, radionuclide distribution was heterogenous after ~ 5 cm downwards. Based on microbial profiles generated, the long-term redox transport and fate of the redox-active radionuclides measured at these sites – Pu and Am – to the depths recorded in this project (~ 30 cm), are not likely to be determined direct interaction with microbially-mediated processes and are more likely influenced by Fe/Mn geochemistry and interaction at these depths.

At the Ravenglass site, the radionuclide activity measured was most abundant in the lower region of the core, at ~ 22 – 26 cm (Figure 4.27), reflecting the historical release of these radionuclides from Sellafield during the 1970s being deposited and Sediments at a rate of ~ 6.8 mm yr^{-1} (Ray *et al*., in prep.). Pu released from Sellafield was mainly present as Fe oxide flocs (Kershaw *et al*., 1992) and it is possible that Pu may be co-associated with Fe surfaces in the effluent (Malcolm *et al*., 1990). As Fe is redox active, Fe cycling in the environment may result in scavenging *via* adsorption or co-precipitation of Pu, or the release of Pu upon microbially-mediated Fe reduction (Malcolm *et al*., 1990). The Pu peaks observed at ~ 22 – 26 cm in the Ravenglass core coincide with low concentrations of porewater Fe and Mn concentrations (Figure 4.25), while solid phase Fe and Mn decrease between 22 and 24 cm (Figure 4.26). This inverse relationship between Fe/Mn concentrations and Pu accumulation suggests that Fe and Mn cycling has little control on Pu distribution in the sediment. At the Mud Patch, there was also no relationship between porewater Fe/Mn profiles and Pu concentration. A previous study showed that Pu in marine particulates exist as insoluble hydroxides, likely in association with Fe/Mn hydroxide colloids (Aston and Stanners, 1981), though the evidence presented here for both Ravenglass and Mud Patch depth cores suggests there is little evidence that shows trace elements controlling the distribution and mobility of Pu.

Since the Pu and Am in these depth cores is independent of Fe and Mn, it is probable that Pu is in a highly refractory residual form (Kimber *et al*., 2015). While Am can be in a mix of phases, its profile in these cores suggests it is possibly also in a residual form similar to Pu (Kimber *et al*., 2015), and that Sellafield-derived Am is poorly soluble in seawater and associated with the same solid sediment component as Pu (Hetherington *et al*., 1976). Since redox conditions in either core at the peak of

Pu/Am accumulation is inferred by the microbial profiles presented here to be sub-oxic, with potential Fe bioreduction, there is unlikely to be any direct bioreduction of Pu, or the production of sulfide from microbial sulfate reduction that may indirectly interact with Pu/Am in this site. The only evidence for microbially-mediated sulfate reduction in either of these sites is in the 0 – 4 cm region of the Ravenglass core, where Pu/Am concentrations are lowest. Furthermore, since it is actinide activity that has been measured here rather than actual concentration, Pu/Am concentration levels are likely to be low, further reducing the likelihood of direct microbial redox interactions.

With regards to Cs activity profiles, which mirrored those of Am and Pu, reflecting historical release from Sellafield, likely reflects Cs being fixed by sediment clay minerals, *e.g.*, illite and vermiculite (Hird *et al*., 1996). The presence of Cs at depths of > 30 cm in the Mud Patch core, and therefore likely pre-1970 peak Sellafield activity, suggests the potential of Cs remobilization from clay minerals at these depths, either *via* cation exchange with ionic species such as K^+ (Sawhney, 1972; Ray *et al*., in prep.).

4.5 References

Al-Qasmi H, Law GTW, Fi LK, Livens FR (2016) Origin of artificial radionuclides in soil and sediment from North Wales. Journal of Environmental Radioactivity, 151(1): 244–249

Al-Qasmi H, Law GTW, Fifield KL, Howe J, Brand T, Cowie GL, Law KA, Livens FR (2017) Deposition of Artificial Radionuclides in Sediments of Loch Etive, Scotland. Journal of Environmental Activity, 187: 45–52

Asshauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomics 16S rRNA data. Bioinformatics, 31(7): 2882–2884

Aston S, Stanners DA (1981) Americium in intertidal sediments from the coastal environs of Windscale. Marine Pollution Bulletin, 12(5): 149–153

Boukhalfa H, Icopini GA, Reilly SD, Neu MP (2007) Plutonium(IV) Reduction by the Metal-Reducing Bacteria *Geobacter metallireducens* GS15 and *Shewanella oneidensis* MR1. Applied and Environmental Microbiology, 73(18): 5897–5903

Brookshaw DR, Pattrick RAD, Lloyd JR, Vaughan DJ (2012) Microbial effects on mineral-radionuclide interactions and radionuclide solid-phase capture processes. Mineralogical Magazine, 76(3): 777–806

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Pena A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7(5): 335-336

Carr A, Blackley MWL (1986) Implications of Sedimentological and Hydrological Processes on the Distribution of Radionuclides the Example of a Salt Marsh near Ravenglass: 529–543

Cleveland CC, Nemergut DR, Schmidt SK, Townsend AR (2007) Increase in soil respiration following labile carbon additions linked to rapid shifts in soil microbial community composition. Biogeochemistry, 89: 229–240

Du Z-J, Wang Z-J, Zhao J-X, Chen G-J (2016) *Woeseia oceani* gen. nov., sp. nov., a chemoheterotrophic member of the order *Chromatiales*, and proposal of *Woeseaiceae* fam. nov. International Journal of Systematic and Evolutionary Microbiology, 66: 107–112

Edgar RC (2017) Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. PeerJ, 5: e3889. https://doi.org10.7717/peerj.3889

Evans PN, Parks DH, Chadwick GL, Robbins SJ, Orphan VJ, Golding SD, Tyson GW (2015) Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. Science, 350(6259): 434–438

Francis AJ (2007) Microbial mobilization and immobilization of plutonium. Journal of Alloys and Compounds, 445-445: 500–505

Gray, J., Jones, S.R., Smith, A.D., 1995. Discharges to the environment from the Sellafield Site, 1951-1992. Journal of Radiological Protection, 15: 99–131.

Gray ND, Head IM (2005) Minerals, mats, pearls and veils: themes and variations in giant sulfur bacteria, in: Gadd GM, Semple KT, Lappin-Scott H (Ed.), Micro-organisms and Earth Systems: 35–71

Halcrow Group (2013) Ravenglass Estuary Complex Sefton Council.

Hetherington JA, Jefferies DF, Mitchell NT, Pentreath RJ, Woodhead DS (1976) Transuranium nuclides in the environment : proceedings of the Symposium on Transuranium Nuclides in the Environment. IAEA

Hird AB, Rimmer DL, Livens FR (1996) Factors affecting the sorption and fixation of caesium in acid organic soil. European journal of Soil Science, 47: 97–104

Huang X-F, Liu YJ, Dong J-D, Qu L-Y, Zhang Y-Y, Wang F-Z, Tian X-P, Zhang S (2014) *Mangrovibacterium diazotrophicum* gen. nov., sp. nov., a nitrogen-fixing bacterium isolated from a mangrove sediment, and proposal of *Prolixibacteraceae* fam. nov. International Journal of Systematic and Evolutionary Microbiology, 64: 875–881

Hug LA, Thomas BC, Sharon I, Brown CT, Sharma R, Hettich RL, Wilkins MJ, Williams KH, Singh A, Banfield JF (2016) Critical biogeochemical functions in the subsurface are associatedwith bacteria from new phyla and little studied lineages. Envrionmental Microbiology, 18(1): 159–173

Kawaichi S, Ito N, Kamikawa R, Sugawara T, Yoshida T, Sako Y (2013) *Ardenticatena maritima* gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum '*Chloroflexi*' isolated from an iron-rich coastal hydrothermal field, and description of *Ardenticatenia* classis nov. International Jounral of Systematic and Evolutionary Microbiology, 63: 2992–3002

Keogh SM, Aldahan A, Possnert G, Finegan P, León Vintró L, Mitchell PI (2007) Trends in the spatial and temporal distribution of 129I and 99Tc in coastal waters surrounding Ireland using Fucus vesiculosus as a bio-indicator. Journal of Environmental Radioactivity, 95: 23–38. doi:10.1016/j.jenvrad.2007.01.009

Kershaw PJ, Swift DJ, Pentreath RJ, Lovett MB (1983) Plutonium redistribution by biological activity in Irish Sea sediments. Nature, 306: 774–775. doi:10.1038/306774A0

Kershaw PJ, Woodhead DS, Malcolm SJ, Allington DJ, Lovett MB (1990) A Sediment History of Sellafield Discharges. Journal of Environmental Radioactivity, 12: 201–241

Kershaw PJ, Pentreath RJ, Woodhead DS, Hunt GJ (1992) A review of radioactivity in the Irish Sea- a report prepared for the Marine Pollution Monitoring Group. Aquatic Environment Monitoring Report.

Kim Y-O, park S, Nam B-H, Jung Y-T, Kim D-G, Bae KS, Yoon J-H (2014) Description of *Lutimonas halocynthiae* sp. nov., isolated from a golden sea squirt (*Halocynthia aurantium*), reclassification of *Aestuariicola saemankumensis* as *Lutimonas saemankumensis* comb. nov. and emended description of the genus *Lutimonas*. International Journal of Systematic and Evolutionary Microbiology, 64: 1984–1990

Kimber RL, Corkhill CL, Amos S, Livens FR, Lloyd JR (2015) Geochemical association of Pu and Am in selected host-phases of contaminated soils from the UK and their susceptibility to chemical and microbiological leaching. Journal of Environmental Radioactivity, 142: 96–102

Kuever J (2014) The Family *Desulfobulbaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (eds) The Prokaryotes. Springer, Berlin, Heidelberg

Leschine SB, Holwell K, Canale-Parola E (1988) Nitrogen Fixation by Anaerobic Cellulolytic Bacteria. Science, 242: 11571159

Lloyd JR (2003) Microbial reduction of metals and radionuclides. FEMS Microbiology Reviews, 27(2-3): 411–425

Mackenzie AB, Scott RD (1993) Sellafield waste radionuclides in Irish Sea intertidal and salt marsh sediments 15.

Mackenzie AB, Cook GT, McDonald P (1999) Radionuclide distributions and particle size associations in Irish Sea surface sediments: implications for actinide dispersion. Journal of Environmental Activity, 44: 275–296\

Malcolm SJ, Kershaw PJ, Cromar NJ, Botham L (1990) Iron and manganese geochemistry and the distribution of ^{239,240}Pu and ^{241}Am in the sediments of the North East Irish Sea. The Science of the Total Environment, 95: 69–87

Marsden OJ, Abrahamsen L, Bryan ND, Philip Day J, Keith Fifield L, Gent C, Goodall PS, Morris K, Livens FR (2006) Transport and accumulation of actinide elements in the near-shore environment: field and modelling studies. Sedimentology, 53: 237–248

McBride MJ (2014) The Family *Flavobaceriaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (Eds) The Prokaryotes. Springer, Berlin, Heidelberg

McCartney M, Kershaw PJ, Woodhead DS, Denoon DC (1994) Artificial radionuclides in the surface sediments of the Irish Sea, 1968–1988. Science of the Total Environment, 141: 103–138. doi:10.1016/0048-9697(94)90022-1

McIlroy SJ, Nielsen PH (2014) The Family *Saprospiraceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (Eds) The Prokaryotes. Springer, Berlin, Heidelberg

Mohr KI, Garcia RO, Gerth K, Irschik H, Müller R (2012) *Sandaracinus amylolyticus* gen. nov., sp. nov., a start-degrading soil myxobacterium, and description of *Sandaracinaceae* fam. nov. International Journal of Systematic and Evolutionary Microbiology, 62: 1191–1198

Morton JT, Toran L, Edlund A, Metcalf JL, Lauber C, Knight R (2017) Uncovering the Horseshoe Effect in Microbial Analyses. mSystems, 2: https://doi.org/ 10.1128/mSystems.00166-16.

Morse JW, Choppin GR (1991) The Chemistry of Transuranic Elements in Natural Waters. Reviews in Aquatic Science, 4: 1–22

Mussmann M, Pjevac P, Krüger K, Dyksma S (2017) Genomic repertoire of the *Woeseiaceae*/JTB255 cosmopolitan and abundant core members of microbial communities in marine sediments. The ISME Journal, 11: 1276–1281

Nelson DM, Lovett MB (1978) Oxidation state of plutonium in the Irish Sea. Nature, 276: 599–601

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Renshaw JC, Law N, Geissler A, Livens FR, Lloyd JR (2009) Impact of the Fe(III)-reducing bacteria *Geobacter sulfurreducens* and *Shewanella oneidensis* on the speciation of plutonium. Biogeochemistry, 94: 191–196

Renshaw JC, Butchins LJC, Livens FR, May I, Charnock JM, Lloyd JR (2005) Bioreduction of Uranium: Environmental Implications of a Pentavalent Intermediate. Environmental Science and Technology, 39(15): 5657–5660

Rusin PA, Quintana L, Brainard JR, Strietelmeier BA, Talt CD, Ekberg SC, Palmer PD, Newton TW, Clark DL (1994) Solubilization of Plutonium Hydrous Oxide by Iron-Reducing Bacteria. Environmental Science and Technology, 28: 1686–1690

Sawhney B (1972) Selective sorption and fixation of cations by clay minerals: a review. Clays and Clay Minerals, 20: 93–100

Schreiber L, Holler T, Knittel K, Meyerdierks A, Amann R (2010) Identification of the dominant sulfate-reducing bacterial partner of anaerobic methanotrohps of the ANME-2 clade. Environmental Microbiology, 12(8): 2327–2340

Shivaji S, Srinivas TNR, Reddy GSN (2014) The Family *Planococcaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (eds) The Prokaryotes. Springer, Berlin, Heidelberg: 303351

Skennerton CT, Chourey K, Iyer R, Hettich RL, Tyson GW, Orphan VJ (2017) Methane-fueled syntrophy through extracellular electron transfer: uncovering the genomic traits conserved within diverse bacterial partners of anaerobic methanotrophic archaea. mBio 8:e00530-17. https://doi.org/10.1128/mBio.00530-17

Srinivas A, Divyasree B, Sasikala Ch, Tushar L, Bharti D, Ramana Ch V (2016) Description of *Jeotgalibacillus alkaliphilus* sp. nov., isolated from a solar salt pan, and *Jeotgalibacillus terrae* sp. nov., a name to replace '*Jeotgalibacillus soli*' Chen *et al.* 2010. International Journal of Systematic and Evolutionary Microbiology 66: 5167-5172, doi: 10.1099/ijsem.0.001491

Takeuchi M, Katayama T, Yamagishi T, Hanada S, Tamaki H, Kamagata Y, Oshima K, hattori M, Marumo K, Nedachi M, Maeda H, Suwa Y, Sakata S (2014) *Methyloceanibacter caenitepidi* ge. no., sp. nov., a facultatively methylotrophic bacterium isolated from marine sediments near a hydrothermal vent. International Journal of Systematic and Evolutionary Microbiology, 64: 462–468

Thauer RK, Jungermann K, Decker K (1977) Energy Conservation in Chemotrophic Anaerobic Bacteria. Bacteriological Reviews, 41: 100–180. doi:10.1073/pnas.0803850105

Timmers PHA, Suarez-Zuluaga DA, van Rossem M, Diender M, Stams AJM, Plugge CM (2016) Anaerobic oxidation of methane associated with sulfate reduction in a natural freshwater gas source. The ISME Journal, 10: 1400–1412

Walker CB, de la Torre JR, Klotz MG, Urakawa H, Pinel N, Arp DJ, Brochier-Armanet C, Chain PSG, Chan PP, Gollabgir A, Hemp J, Hügler M, Karr EA, Könneke M, Shin M, Lawton TJ, Lowe T, Martens-Habbena W, Sayavedra-Soto LA, Lang D, Sievert SM, Rosenzweig AC, Manning G, Stahl DA (2010) *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in

globally distributed marine crenarchaea. Proceedings of the National Academy of Sciences of the United States of America, 107(19): 8818–8823

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y (2006) *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum *Chloroflexi*. International Journal of Systematic and Evolutionary Microbiology 56: 1331-1340, doi: 10.1099/ijs.0.64169-0

Yang S-J, Choo Y-J, Cho J-C (2007) *Lutimonas vermicola* gen. nov., sp nov., a member of the family *Flavobacateriaceae* isolated from the marine polychaete *Periserulla leucophryna*. International Hournal of Systematic and Evolutionary Microbiology, 57: 1679–1684

Chapter 5. Using Next-Generation Sequencing Techniques to Investigate Microbial Diversity and Function in Sediment Cores from Loch Etive, Scotland

5.1 Introduction

Loch Etive, which has been studied extensively from a geochemical perspective (see below), is a loch 5 km north of the town of Oban, on the west coast of Scotland and has been well characterised from a hydrography and geochemical perspective. The loch is a glacially overdeepened valley 31.6 km long and between 1.2 – 1.6 km wide, giving a small surface area of ~ 26 km^2 (Wood *et al.*, 1973). The loch is comprised of two distinct basins separated by a sill in the middle of the loch at a site named Bonawe (Edwards and Edelsten, 1977) (Figure 1.7).

The outer basin that connects to the ocean at Connel runs east-west in direction, is mostly estuarine, and has a maximum depth of 60 m. Fresh water from the River Awe enters the outer basin at the western-most point of the loch, at a widening of the basin called Airds Bay. The water in the outer basin is generally well-mixed and oxygenated. Macrofauna such as worms and bivalves are abundant in the upper sediments, below which macrofaunal activity is considered absent due to anaerobic conditions (Ridgway and Price, 1987). Most organic matter environments similar to the loch is mineralised in the water column and in suspended sediments (De Hass *et al.*, 2002) with only small amounts reaching the seafloor (Libes, 1992). Much of the organic carbon that does reach the floor of the loch is recalcitrant lignin (Coûteaux *et al.*, 1991), which may be used as a biomarker as it is generally resistant to biodegradation (Loh *et al.*, 2008). Sedimentation rates of the outer basin are around ~ 0.6 cm per year (Al-Qasmi *et al.*, 2018).

The inner basin runs north-south at an almost-90° angle to the outer basin. The inner basin is much deeper, with a maximum depth of 150 – 170 m, and is generally less mixed and oxygen-poor, and generally considered to be anaerobic at depth. The River Etive enters the inner basin at the northern-most tip. Loch Etive is characterised by a lack of interaction between these two basins due to the shallow sill at Bonawe, except for periodic flushing of the inner basin with fresher, oxygenated waters, caused by overflowing of the sill due to spring tides, regenerating the deeper basin water around every 16 months (Edwards and Edelsten, 1977; Ridgway and Price, 1987). Sedimentation rates of the inner basin are slower than that of the outer basin, at ~ 0.3 cm per year (Al-Qasmi *et al.*, 2018).

The presence of heavy metal contaminants, likely from an anthropogenic source, has been identified in sediments collected from Loch Etive (Ridgway and Price, 1987). These metals, including Zn, Pb, and ^{137}Cs of which the later probably originated from the Sellafield nuclear plant in Cumbria (England), were found to be enriched in sediments from Loch Etive, with higher values found in the deeper, inner basin, likely due to less biomixing. In the context of the subject focus of this thesis, a recent study showed that a range of natural ($^{234, 238}$U) and artificial (^{137}Cs, ^{241}Am, ^{238}Pu, $^{239/240}$Pu, and ^{236}U) radionuclides from Sellafield have been deposited in the sediments of Loch Etive (Al-Qasmi *et al.*, 2018).

The contrasting conditions of the two basins, despite being of the same loch, and the well-characterised hydrography and geochemistry of the site, made it an ideal location for a natural laboratory to investigate radionuclide deposition from Sellafield at a greater distance than Ravenglass. However, the microbial community of the loch has, hitherto, not been investigated using contemporary molecular biological or high-throughput DNA sequencing techniques.

The site was initially chosen as a 'control' site for the entire project, as it was believed to be a site that had negligible levels of redox-active radionuclide contamination (*i.e.*, not greater than background levels) from either geogenic or anthropogenic sources, and additionally, as a sediment system with redox conditions that followed the normally-expected or 'classical', thermodynamically-controlled vertical stratification of microbial redox processes with depth (Thauer *et al.*, 1977), due to the constant sedimentation and lack of mixing observed in the inner basin. In light of the publication of the radionuclide study by Al-Qasmi *et al.* (2018), the focus of this particular study shifted from that of treating the loch as a radionuclide control, to one of being a redox control – where we could directly compare the expected transitions of microbial community taxonomic and functional assemblages with increasing depth and related redox succession, with that observed at other sites in the project. It was envisaged that creating microbial taxonomic and functional profiles associated with sediment cores collected from both the outer and inner basin of the lochs could be used to explain and infer observed geochemical conditions at the site and be used to anticipate the long-term fate of nutrients, metals, and radionuclide contaminants in the loch. The presence and abundance of key functional microbial taxa and genes can be directly linked to certain biogeochemical cycles. This study used a range of DNA-based molecular biological tools, and subsequent bioinformatics-based analysis, to create taxonomic and functional profiles associated

with sediment cores collected from the outer and inner basin. These profiles were used to create inferences about the geochemical conditions associated with the sediment cores. These inferences were compared with directly measured geochemical analyses including XRF and ICP-MS.

5.2 Methods

5.2.1 Sample Collection

One depth core was collected each from two separate parts of Loch Etive (56°26'90" N, 5°14'77" W) near the town of Oban (Scotland, UK) in June 2017. One core was collected from the outer basin at Airds Bay of Loch Etive by Bonawe, from a depth of 71 metres (referred to hereafter as the Shallow Basin). The other core was collected from the inner basin of Loch Etive, from a depth of 147 metres (referred to hereafter as the Deep Basin). See Chapter 2.1.3.

5.2.2 DNA Extraction and Quantification

DNA was extracted from ~ 500 mg of sediment from each sample *via* FastDNA Spin Kit for Soil as per manufacturer's instructions (MP Biomedicals, USA). For all analysis, samples within each core were separated into arbitrary groups corresponding to depth based on the resolution they were sampled at, to allow for statistical comparison of values between sample groups and identify patterns of abundance associated with depth. Sample groups were comprised of: 0.5 – 2 cm, at 0.5 cm resolution for Group 1; 3 – 10 cm at 1 cm resolution for Group 2; 12 – 20/22 cm at 2 cm resolution, for the Shallow and Deep Basins respectively, and; 25/27 – 42/47 cm at 5 cm resolution for Shallow and Deep Basin respectively (including the final sample at 7 cm resolution for the Shallow Basin). See Chapter 2.2.

5.2.3 'End Point' PCR Functional Gene Series

The primers and PCR protocols used are shown in Table 3.1. The 'end point' PCRs were used to give a qualitative indication of target gene abundance within the depth cores and test the utility of this simple approach to informing redox conditions. See Chapter 2.3.

5.2.4 16S rRNA Gene Sequencing Preparation

All samples from both cores (including blanks) were prepared for 16S rRNA gene sequencing by undergoing PCR using the 515F and 926R primers (*Escherichia coli* numbering) targeting the V4–V5 region of the 16S rRNA gene. See Chapter 2.5.1.

5.2.5 16S rRNA Gene Sequencing Pipeline Analysis

Multiplexed Ion Torrent FASTQ files were imported into and analysed using QIIME2 (Caporaso *et al.*, 2011). See Chapter 2.6.

5.2.6 Tax4Fun Analysis

The SILVA119 closed-reference OTU tables and taxonomic assignments derived from 16S rRNA gene sequencing analysis of both cores were run on the Tax4Fun R package (Asshauer *et al.*, 2015) to predict KEGG Orthologs. KEGG Orthologs from catabolic pathways of interest were extracted and are included in Table 2.4. See Chapter 2.6.9.

5.2.7 Geochemical Analysis

Geochemical analysis was performed at the University of Manchester. See Chapter 2.9.

5.2.8 Environmental Fitting

Detrended correspondence analysis (DCA) plots were produced using the 'vegan' package in R 3.5.4. Additionally, geochemical data was superimposed onto the DCA using environmental fitting. See Chapter 2.8.2.

5.3 Results

5.3.1 DNA Extractions

The volume of DNA extracted from the Shallow Basin depth core decreased significantly with increasing depth (Appendix E.1A), ranging from a peak of 4.22

ng/μL at 3 cm to 0.46 ng/μL at 42 cm (Figure 5.1A). DNA volumes extracted from the Deep Basin depth core also decreased significantly with depth (Appendix E.1B), albeit not with the progressive trend observed in the Shallow Basin but in a more spatially-short area and ranged from a peak of 2.99 ng/μL at 5 cm and a low of 0.69 ng/μL at 42 cm (Figure 5.1B). Between 0 – 6 cm, the average volume of DNA extracted was 2.42 ± 0.12 ng/μL, whereas between 7 – 37 cm, this lowered to an average of 1.45 ± 0.06 ng/μL, and this dropped again to 0.69 and 0.88 ng/μL at 42 and 47 cm respectively.

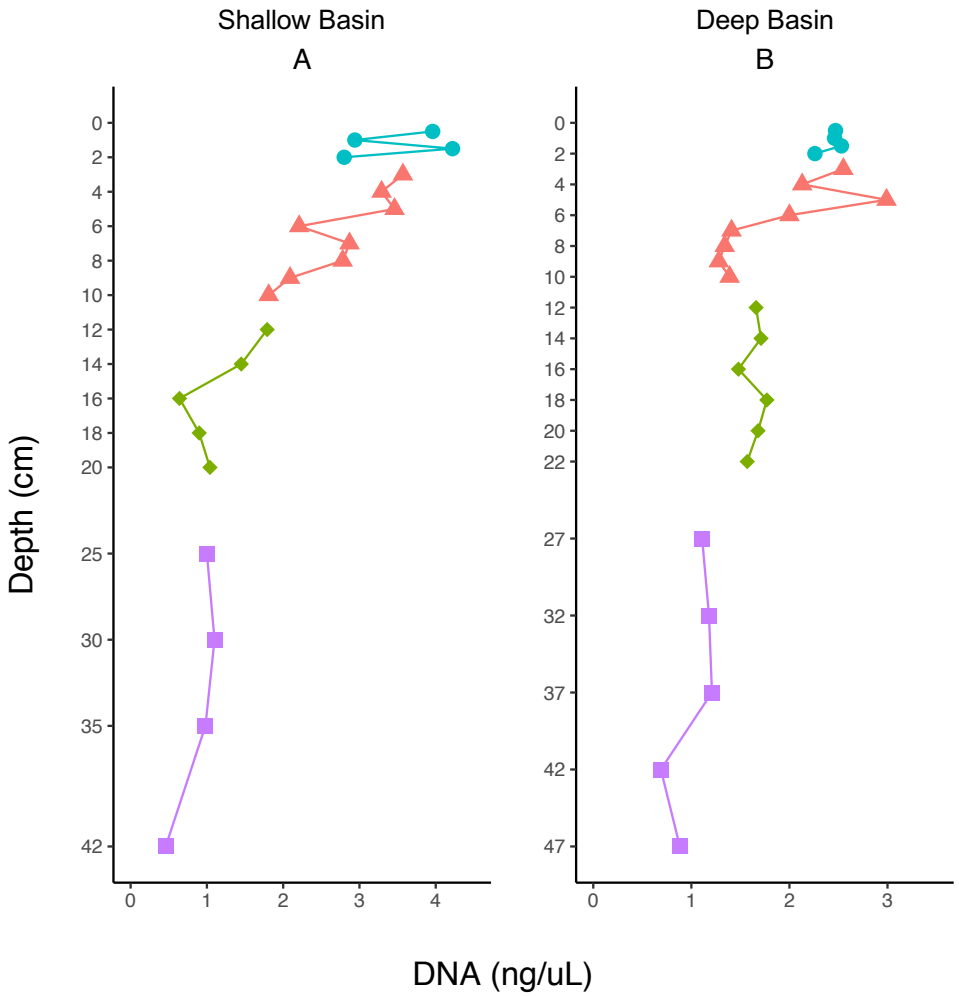


Figure 5.1. Depth profiles showing the amount of DNA in ng/μL extracted from each sample from the Shallow Basin depth core (A). and the Deep Basin depth core (B). Colours/points refer to sample groups as defined by sampling resolution: Blue circles = group 1 (0 – 2 cm, 0.5 cm resolution); red triangles = group 2 (3 – 10 cm, 1 cm resolution); green diamonds = group 3 (12 – 20/22 cm for Shallow and Deep Basin respectively, 2 cm resolution); purple squares = group 4 (25/27 – 42/47 cm for Shallow/Deep Basin respectively, 5 cm resolution, including the final sample at 7 cm resolution from Shallow Basin).

5.3.2 'End Point' PCR Functional Gene Series

'End point' PCR band intensities for functional genes were quantified *via* BioNumerics, and bands given a grey-scale 8-bit value (pixel value) based on the band size, thickness. Band intensities as defined by the pixel value are presented as a qualitative inference of gene abundance within each depth sample, *i.e.*, a higher band intensity denotes a higher gene abundance within a specific sample.

5.3.2.1 *Shallow Basin Depth Core 'End Point' PCR*

The archaeal ammonia monooxygenase (*amoA*) gene was very low throughout the core, never exceeding a band intensity of 36, with an average band intensity of 25.19 ± 1.94 (Figure 5.2A). AOV and Tukey HSD showed only a borderline statistical difference in band intensities between any sample groups (*i.e.*, p value > 0.05) (Appendix E.2A). The bacterial *amoA* gene, in contrast, had a higher abundance in the core and was generally higher in abundance in the upper region of the core, with an average band intensity of 219.92 ± 11.87 between 0 – 10 cm, which lowered to an average of 126.2 ± 14 between 12 – 20 cm and increased again to an average of 189.5 ± 24.61 between 25 – 42 cm (Figure 5.2B). AOV showed a statistical difference in band intensity across all sample groups, and Tukey HSD showed statistical differences in band intensities between sample group pair 2 – 3 (Appendix E.2B). The denitrifying nitrite reductase *nirK* gene was low in abundance similar to the archaeal *amoA*, never exceeding a band intensity of 36 (Figure 5.2C). AOV showed a statistical difference in the *nirK* band intensity across all sample groups, and Tukey HSD showed statistical differences between group pairs 2 – 3, and 3 – 4 (Appendix E.2C). The other denitrifying nitrite reductase gene, *nirS*, was higher in abundance than the *nirK* as well as having a higher band intensity in the upper region of the core having an average of 85.14 ± 11.33 between 0 – 5 cm, which dropped to an average of 49.43 ± 4.58 between 6 – 42 cm (Figure 5.2D). AOV showed a statistical difference in *nirS* band intensity across all sample groups, and Tukey HSD showed a statistical difference between sample pair 1 – 4 (Appendix E.2D). The dissimilatory sulfite reductase *dsrA* gene would not to amplify from these extracts including the DNA positive, so it was not possible to discern whether the *dsrA* gene was present. However, the methanogenic *mcrA* gene was not detected in any samples, where the DNA positive did amplify correctly, indicating a lack of this function.

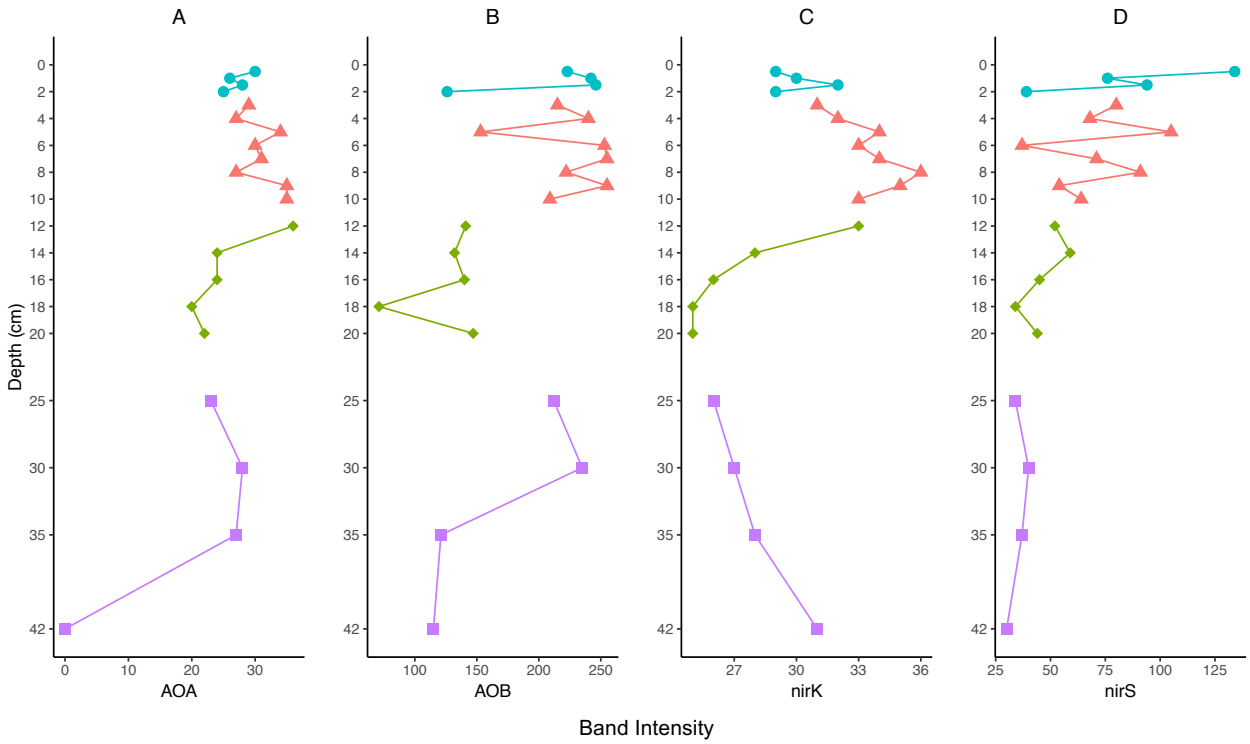


Figure 5.2. Shallow Basin Depth Core: Functional gene 'end point' PCR band intensities for the Ravenglass core. Values on the Y axis represent depth in cm, values on the X axis represent the band intensity of the PCR amplicon as calculated by BioNumerics. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*.

5.3.2.2 *Deep Basin Depth Core 'End Point' PCR*

Similar to the Shallow Basin depth core, the archaeal *amoA* gene was low in abundance throughout the core, with an average band intensity of 42.78 ± 1.02 (Figure 5.3A). AOV and Tukey HSD did not show statistical differences in archaeal *amoA* abundance across all sample groups and between sample groups (Appendix E.3A). The bacterial *amoA* gene had a higher abundance than the archaeal *amoA* but was largely confined to the upper-most part of the core, with an average band intensity of 194 ± 17.79 between 0 – 3 cm, which lowered to an average of 88.33 ± 6.23 between 4 – 6 cm and lowered again to an average of 46 ± 5.32 between 7 – 47 cm (Figure 5.3B). AOV showed a statistical difference in band intensity across all sample groups, and Tukey HSD showed statistical differences between group 1 and all three other groups (Appendix E.3B). The *nirK* gene generally had a low band intensity in the Deep Basin as it did in the Shallow Basin depth core and was most abundant in the very upper-most part of the core, with an average band intensity of 106 ± 12.17 between 0 – 1.5 cm, and 35.15 ± 2.21 between 2 – 47 cm (Figure 5.3C).

AOV showed a statistical difference in the *nirK* band intensities across all sample groups, and Tukey HSD showed differences between group 1 and the three other groups (Appendix E.3C). The *nirS* gene was also low in abundance and apparently restricted to the 0 – 1.5 cm region, where the average band intensity was 44.33 ± 2.19, and 33.1 ± 0.62 between 2 – 47 cm (Figure 5.3D). AOV showed a statistical difference in the *nirS* band intensities across all sample groups, and Tukey HSD showed statistical differences between group 1 and all other three groups (Appendix E.3D).

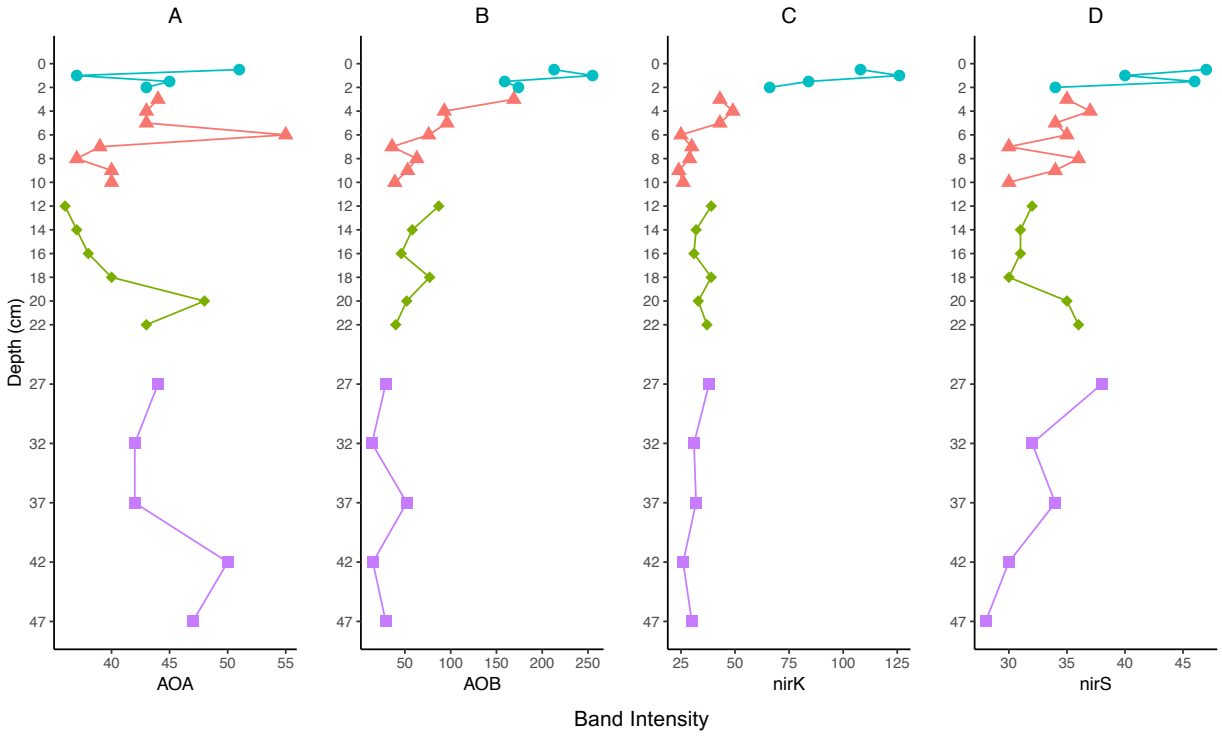


Figure 5.3. Deep Basin Depth Core: Functional gene end point PCR band intensities for the Ravenglass core. Note that the X axes have different ranges for each gene. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*.

5.3.3 16S rRNA Gene Sequencing Results

Sequencing produced 1,817,344 reads in the demultiplexed Shallow Basin depth core library, and 1,511,562 reads in the demultiplexed Deep Basin depth core library. A total of 7,908 ASVs from 1,372,387 reads were identified in the Shallow Basin core, and 9,352 ASVs from 1,156,592 reads were identified in the Deep Basin core by DADA2 during pipeline analysis. ASV tables were rarefied to 21,346 reads per depth sample for the Shallow Basin depth core, and to 30,000 for the Deep Basin depth core (therefore excluding the 18 cm sample), for core phylogenetic metric diversity analysis in QIIME2. ASV taxonomic abundances were investigated at genus

level, with a total of 1,156 ASVs at genus level from the Shallow Basin and 1,161 ASVs at genus level from Mud Patch.

5.3.3.1 *Alpha Diversity Analysis*

Shannon's species richness had no discernible patterns with depth in either core (Figures 5.4A and 5.4B) and there were no statistical differences in index values with depth (Appendix E.4A and 5.5B). Faith's phylogenetic diversity generally increased with depth in both cores (Figures 5.4C and 5.4D). AOV and Tukey HSD did not show statistical differences in index values between sample groups in the Shallow Basin core (Table 5.5B), while the Deep Basin core did have a statistical difference based on AOV and Tukey HSD showed statistical differences between sample group pairs 1 – 4 and 3 – 4 (Appendix E.4D). Pielou's species evenness generally decreased with depth in both cores (Figures 5.4E and 5.4F). AOV showed a statistical difference in Pielou's species evenness index across all sample groups from the Shallow Basin, and Tukey HSD showed statistical differences between sample group pairs 1 – 3, 1 – 4, and 2 – 4 (Appendix E.4E). AOV showed a statistical difference in Pielou's species evenness in the Deep Basin, and Tukey HSD showed a statistical difference between sample group pair 1 – 2 (Appendix E.4F).

5.3.3.2 *Beta Diversity Analysis*

The non-metric multidimensional scaling (nMDS) plot (based on Bray Curtis dissimilarity) and the detrended correspondence analysis (DCA) plot of the ASV microbial community assemblages of the two cores as derived from 16S rRNA gene sequencing revealed two separate community assemblages split by site, with depth samples of both cores distributed along the nMDS axes in the same depth-related pattern (Figure 5.5A) and diverging separately on the DCA plot (Figure 5.5B). The nMDS and DCA plots of the taxonomic classifications of the ASVs collapsed at class level showed greater similarity between the 0.5 – 10 cm samples of both cores compared to the ASV community assemblage, and showed a similar depth-related distribution of samples from both cores in the same order, with the exception of the Shallow Basin depth core which, after ~ 12 cm, was distributed in a 90° angle to the Deep Basin samples (Figures 5.6A and 5.6B).

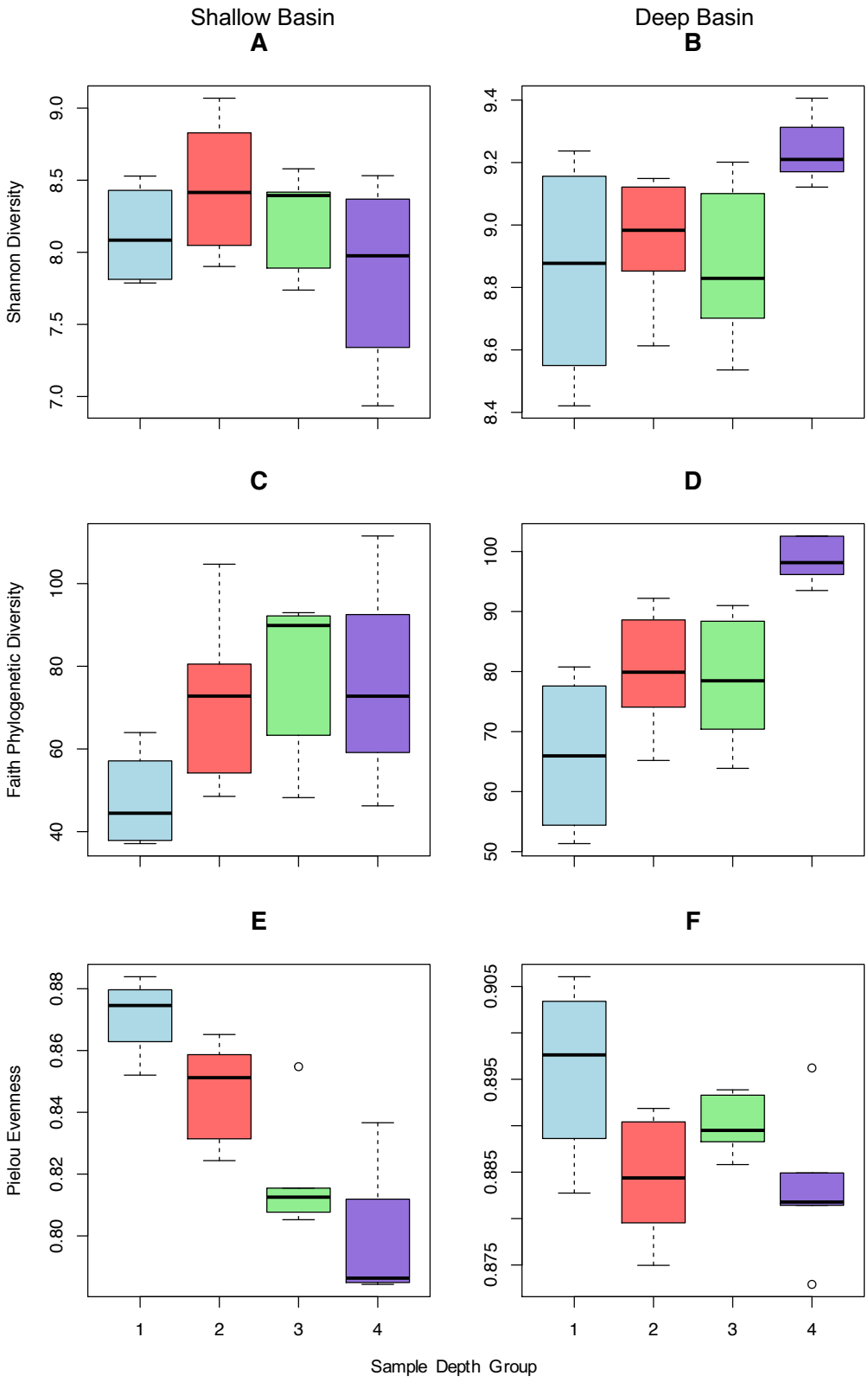


Figure 5.4. Whisker box plots of three alpha diversity indices for the Shallow Basin and Deep Basin depth cores. Boxes denote samples that were sampled at the same resolution: Group 1 = 0.5 cm; 2 = 1 cm; 3 = 2 cm; 4 = 5 cm. (A). Boxplot of Shannon richness diversity of Shallow Basin. (B). Shannon richness diversity of Deep Basin. (C). Faith phylogenetic diversity of Shallow Basin. (D). Faith phylogenetic diversity of Deep Basin. (E). Pielou species evenness of Shallow Basin. (F). Pielou species evenness of Deep Basin.

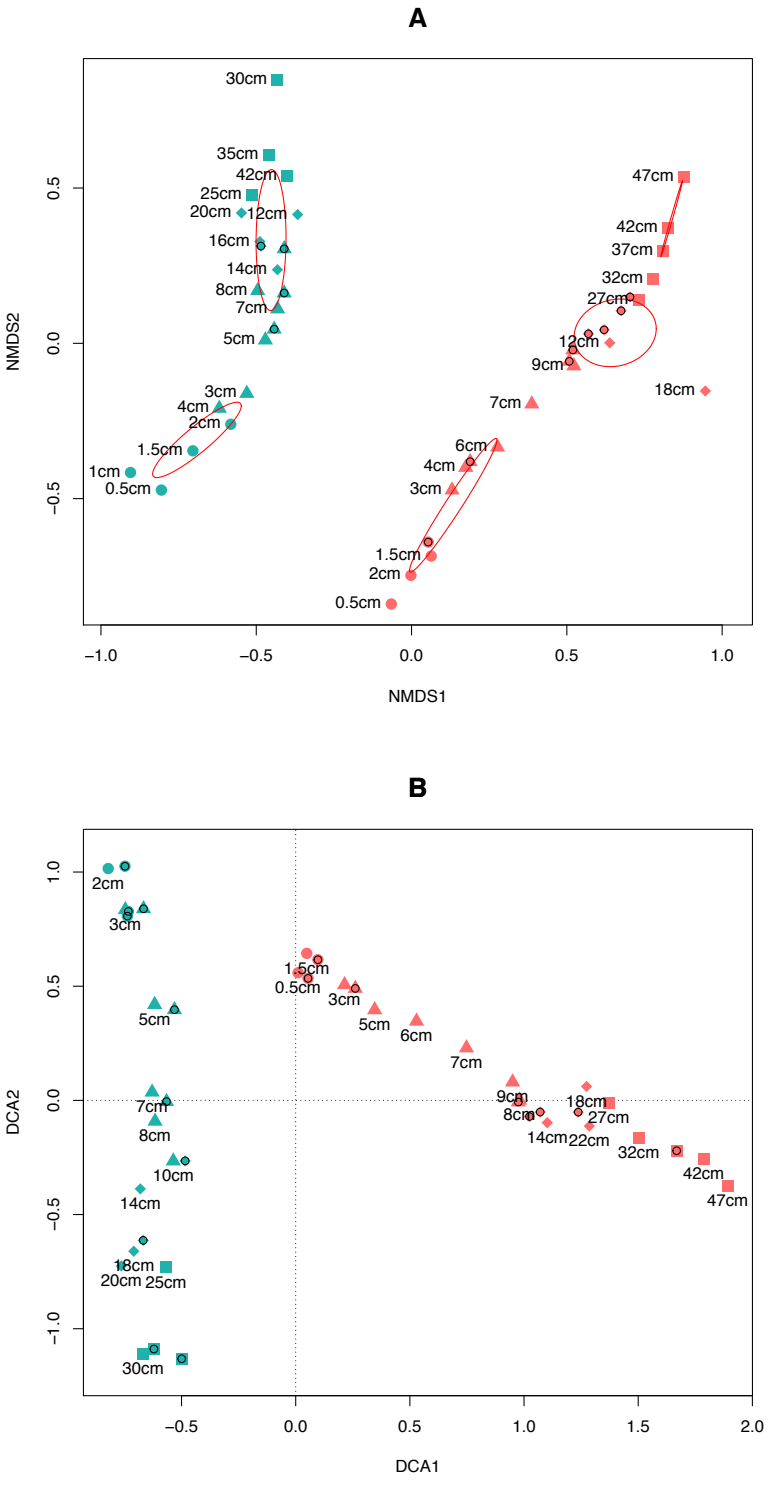


Figure 5.5. (A). Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity metric of samples from both depth cores. Blue denotes Shallow Basin depth core samples and red denotes Deep Basin depth core samples. Point shapes denote sampling depth groups. Black circles denote samples without a label (due to space constraints). Red outline circles denote hierarchical clustering of samples based on centroid dissimilarity. (B). Detrended correspondence analysis (DCA) of samples from both depth cores.

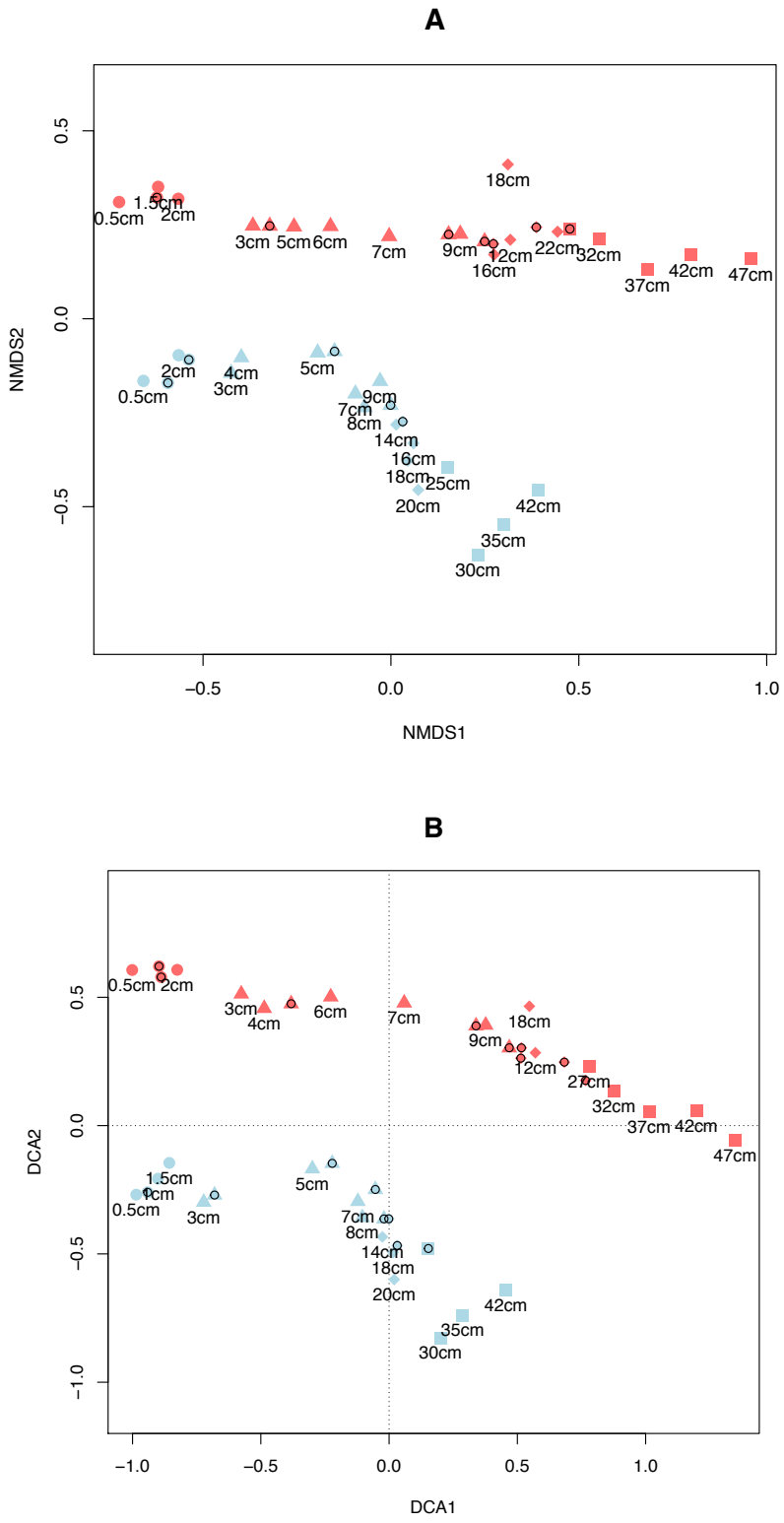


Figure 5.6. Ordination of the ASV taxonomic classification assemblages collapsed at genus level of both depth core samples. Blue samples denote Shallow Basin depth core, and red samples denote Mud Patch depth core. Points refer to sample depth groups. (A). Non-metric multidimensional scaling (nMDS) plot based on Bray Curtis dissimilarity of taxonomy collapsed at genus level; (B). Detrended Correspondence Analysis (DCA) of taxonomy collapsed at genus level.

2D PCoA plots of Shallow Basin samples based on Bray Curtis dissimilarity resulted in a depth-related ‘arch’ distribution of samples along PCs 1 and 2 (Figure 5.7A). This ‘arch’ effect is a result of the restriction of many ASVs to a few sample depths, and therefore many zeroes in the ASV table, defined as a sparse table, and results in a loss of community dissimilarity resolution after a certain point of dissimilarity, known as saturation (Morton *et al.*, 2017a). Weighted UniFrac dissimilarity showed, though less-defined, a similar arching of samples to Bray Curtis (Figure 5.7B). Unweighted UniFrac showed a general distribution of samples with depth (Figure 5.7C). PERMANOVA across all sample groups and pairwise PERMANOVA between sample groups showed a statistical difference in the distances between the sample groups for all pairs of samples across all dissimilarity methods, except group pair 3 – 4 of unweighted UniFrac dissimilarity (Appendix E.5). The nMDS and DCA plots of the Shallow Basin depth core ASV community assemblage showed a general distribution of samples in a depth-related pattern along one axis, with heterogeneity along the second axis between individual samples (Figure 5.8).

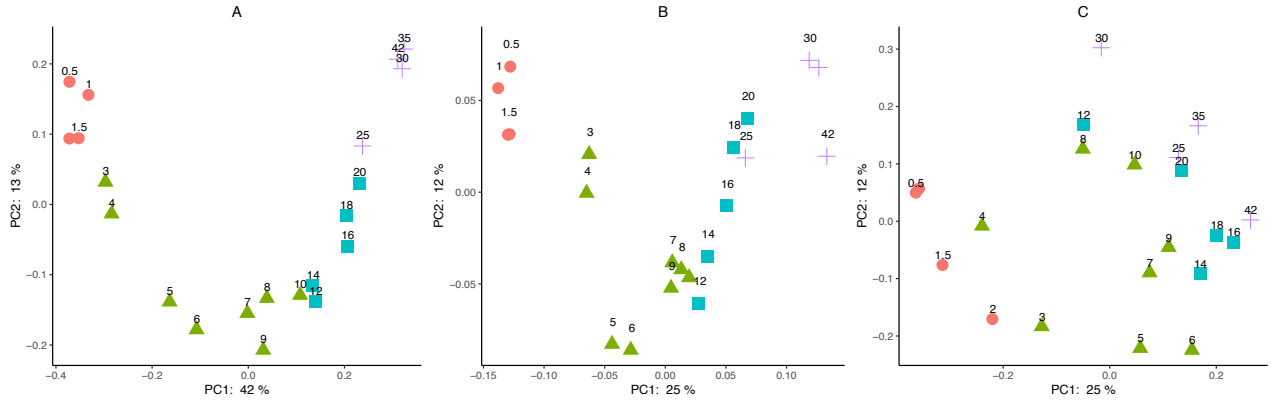


Figure 5.7. Shallow Basin Depth Core: PCoA based on three dissimilatory metrics. Colours/points refer to sampling resolution groups. Red circles = Group 1 (0 – 2 cm); Green triangles = Group 2 (3 – 10 cm); Blue squares = Group 3 (12 – 20 cm); Purple crosses = Group 4 (25 – 42 cm). Numbers represent depth in cm. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac dissimilarity; (C). Unweighted UniFrac dissimilarity.

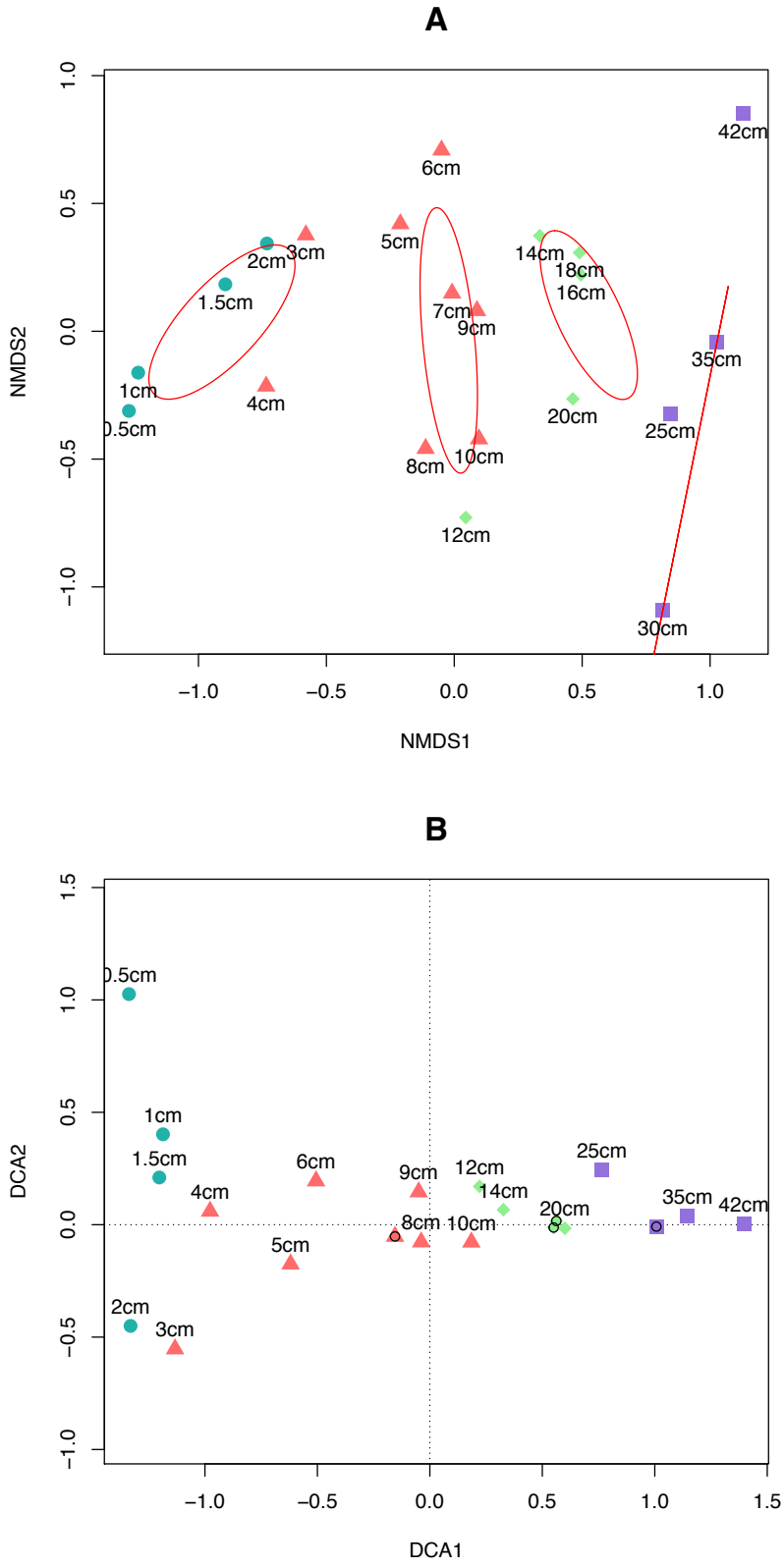


Figure 5.8. Shallow Basin Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups.

2D PCoA plots of Deep Basin samples based on Bray Curtis dissimilarity showed again the 'arch' effect described above, with samples in almost-perfect order of depth (Figure 5.9A). Both weighted and unweighted UniFrac showed similar

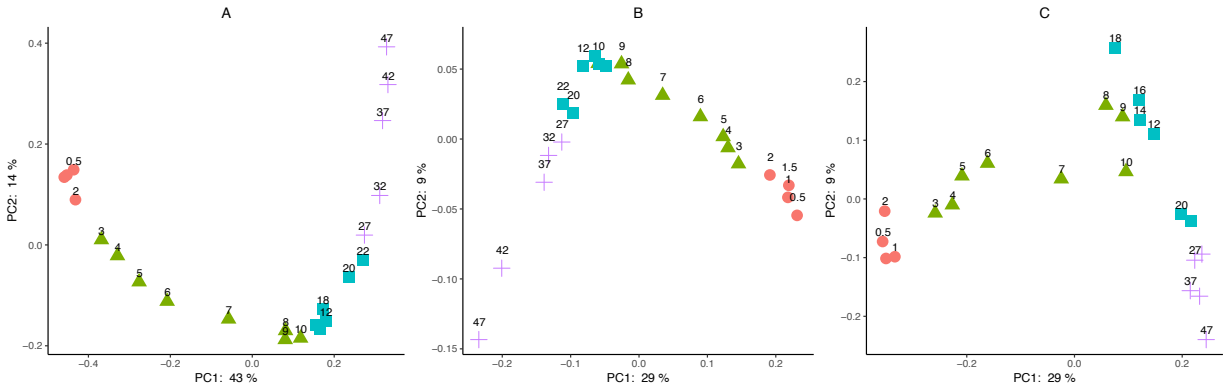


Figure 5.9. Deep Basin Depth Core: PCoA based on three dissimilatory metrics. Colours/points refer to sampling resolution groups. Red circles = Group 1 (0 – 2 cm); Green triangles = Group 2 (3 – 10 cm); Blue squares = Group 3 (12 – 22 cm); Purple crosses = Group 4 (27 – 47 cm). Numbers represent depth in cm. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac dissimilarity; (C). Unweighted UniFrac dissimilarity.

patterns of sample distributions to Bray Curtis (Figures 5.9B and 5.9C). PERMANOVA showed a statistical difference in group distances across and between all sample groups for all three dissimilarity methods (Appendix E.6). The nMDS plot representing the samples from the Deep Basin depth core showed a more defined depth-related distribution of samples along the nMDS1 axis, with four clusters of samples being identified. Sample representing 20 cm was an outlier on the plot (Figure 5.10A). The DCA plot of the Deep Basin ASV community (Figure 5.10B) showed a half-diamond shape emerging after ~ 10 cm, a result of the detrending contorting the ordination space (Hill and Gauch, 1980).

5.3.4 Differential Abundance Analysis

The heatmap showing whether each ASV is statistically differently abundant from other ASVs at in a given sample, *i.e.*, depth, presented with the hierarchical clustering tree of co-occurring ASVs, showed that the Shallow Basin and Deep Basin have their own distinct communities of ASV that are proportionally more abundant in one than the other with some degree of overlap, and that the tree is divided into two main branches, one for each core (Figure 5.11). The boxplot of the log ratio of the '*y*0' balance shows that the taxa comprising the '*y*0' numerator are more abundant

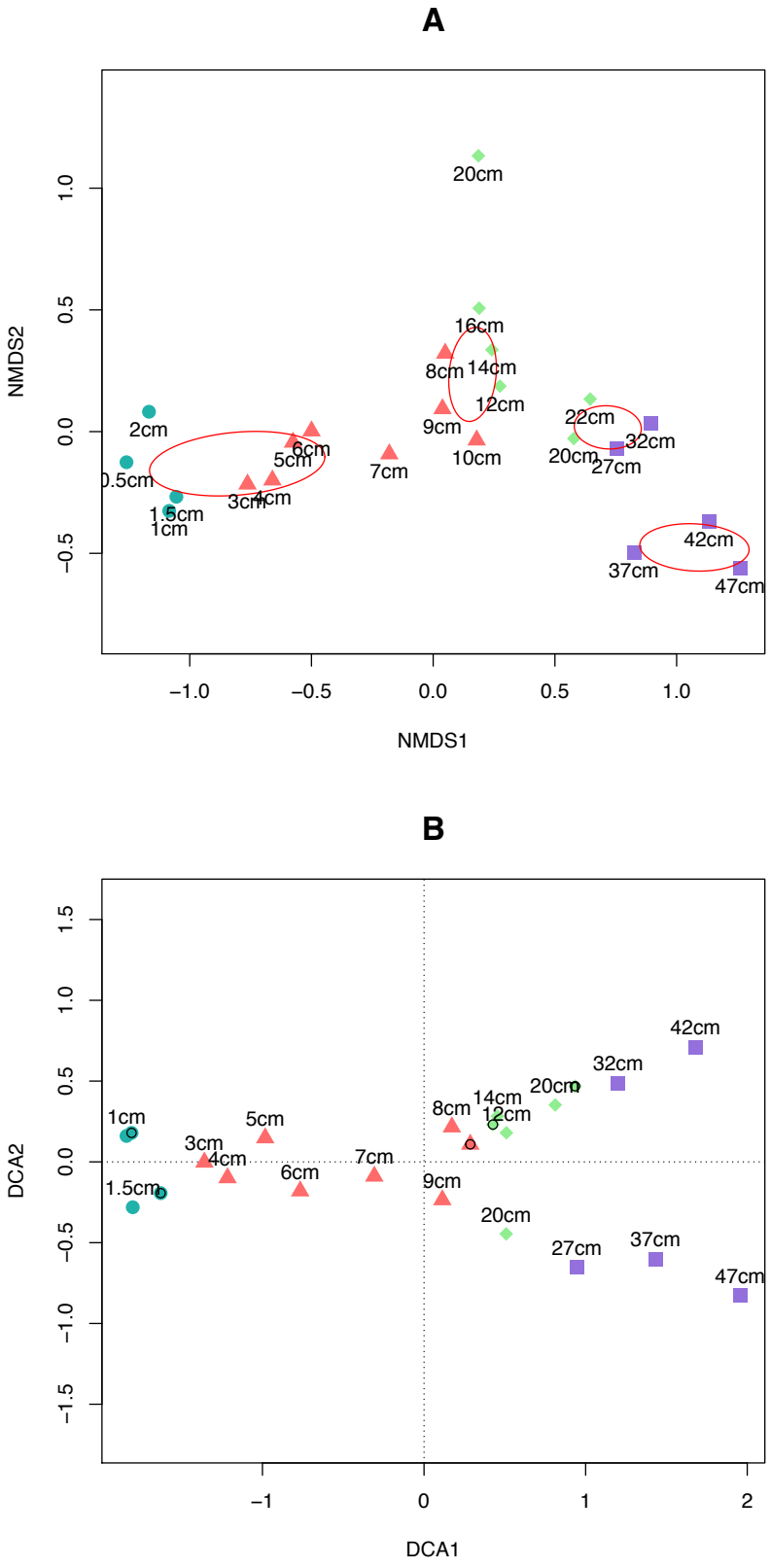


Figure 5.10. Deep Basin Depth Core: (A). nMDS plot based on Bray Curtis dissimilarity of the microbial community assemblages between samples. Red circles denote average hierarchical clustering of samples. (B). DCA plot. Colours and points denote sample depth groups.

than the taxa in the '*y*0' denominator in the Shallow Basin compared to the Deep Basin (Figure 5.12). The proportion plot of '*y*0' balance at order level shows that the notable ASV taxonomies that were more proportionally abundant in the Shallow Basin were three *Clostridiales* ASVs, whereas in Deep Basin it was three *Ignavibacteriales* and a *Myxococcales* ASVs (Figure 5.13).

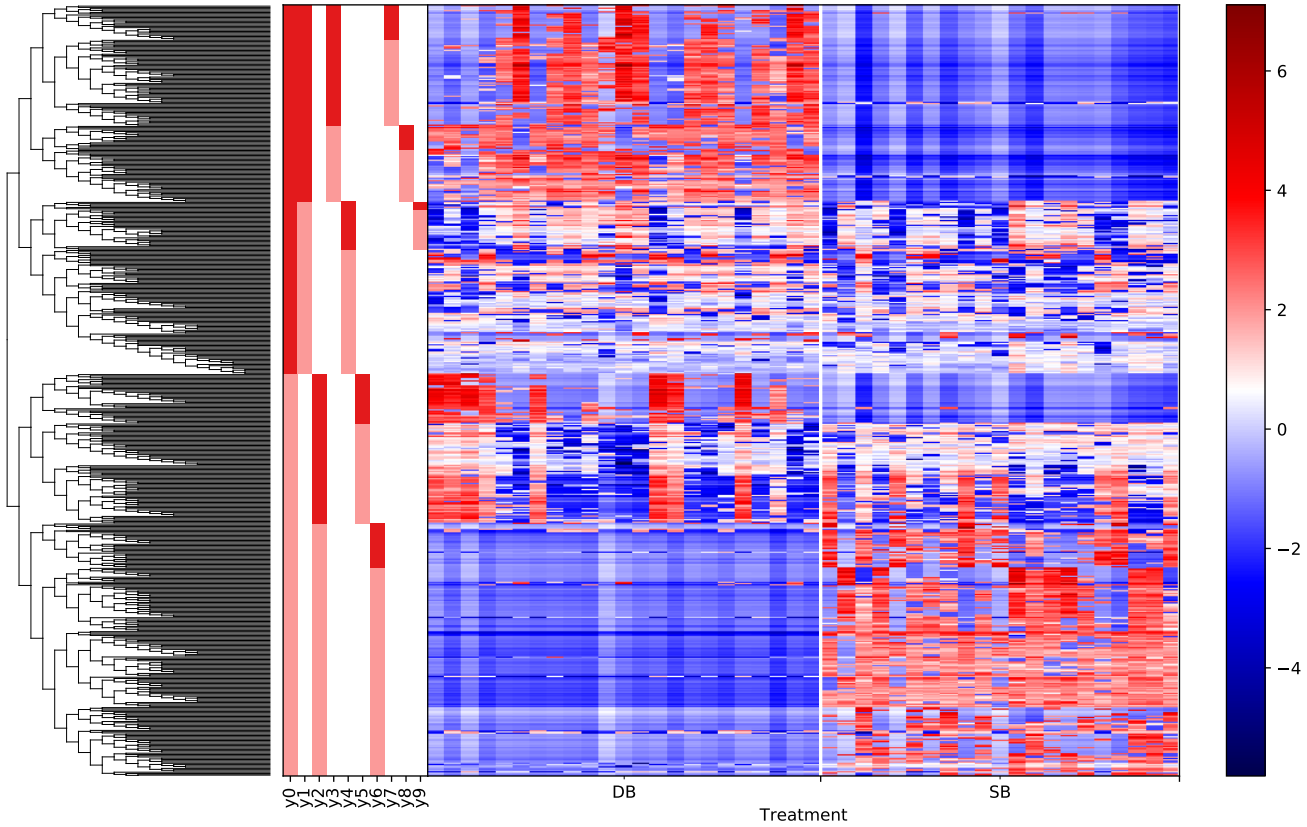


Figure 5.11. Heatmap of individual ASV p-values that have been log transformed and dendrogram of hierarchical clustering of differentially abundant ASVs between the two depth cores. The spectrum of the heatmap indicates that an ASV that is red has a more positive negative-log p-value and thus more proportionally abundant in the respective sample, where as a blue ASV has a more negative negative-log p-value. Dendrogram denotes Ward's hierarchical clustering of ASVs that co-occur with each other. Red bars labelled '*y*' denote the balances, with light red representing the numerator and dark red the denominator. DB = Deep Basin depth core, SB = Shallow Basin depth core.

5.3.5 Taxonomic Analysis

ASV taxonomic classifications were collapsed at genus level, *i.e.*, all ASVs that had the same classification to genus level were merged into one ASV. Not all ASVs classified down to a specific named genus however, and therefore the lowest specific named classification level is presented, but still represents ASVs merged at the genus level.

5.3.5.1 *Shallow Basin Depth Core Taxonomy*

Firstly, it was observed that the split between bacterial and archaeal abundances in the sample depths ranged from 99.43% bacterial and 0.57% archaeal

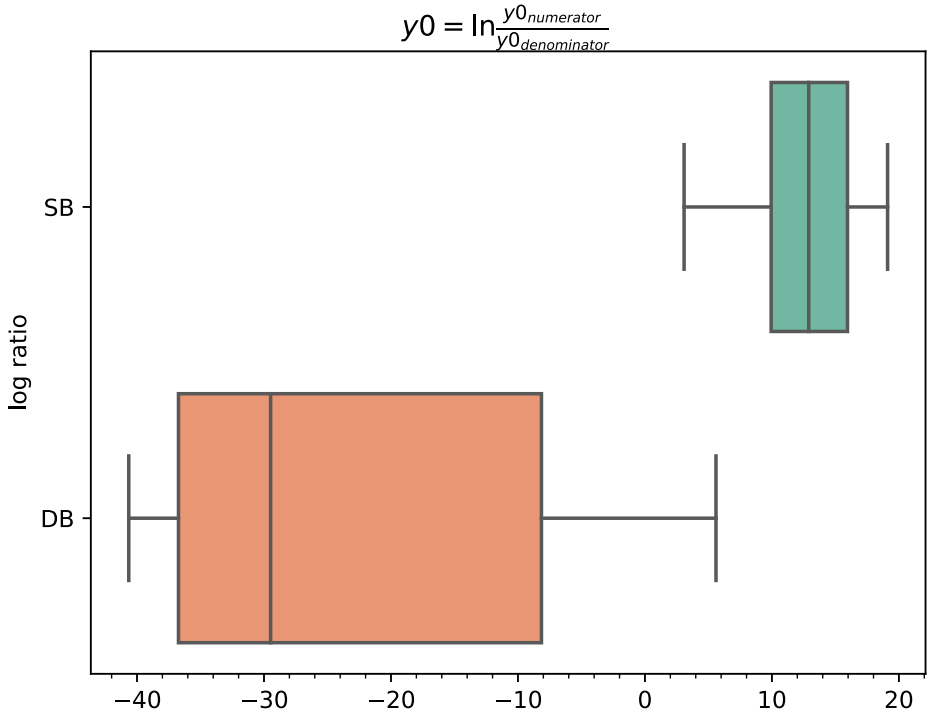


Figure 5.12. Boxplot of the log ratio of the abundance of the taxa in '*y*0' numerator (light red bar in Figure 5.11) and 'y0' denominator (dark red bar in Figure 5.11) from the Shallow Basin (top, green) and Deep Basin (bottom, orange).

at 1 cm to 94.33% bacterial and 5.22% archaeal (remainder unassigned) at 5 cm. The ten most abundant taxonomic groups when ASVs were collapsed to genus level are presented in Figure 5.14 and here in order of descending total abundance. These ten taxa accounted for between 34% (6 cm) and 48% (30 cm) of the entire community assemblages. Of this ten, four were from the order *Clostridiales*. The most abundant *Clostridiales* taxon was from the genus *Tepidibacter* spp., which increased in abundance with depth, ranging from 2.6% at 0.5 cm to 20.4% at 30 cm. AOV showed a statistical difference in *Tepidibacter* spp. abundance across all sample groups, with Tukey HSD showing statistical differences in abundances

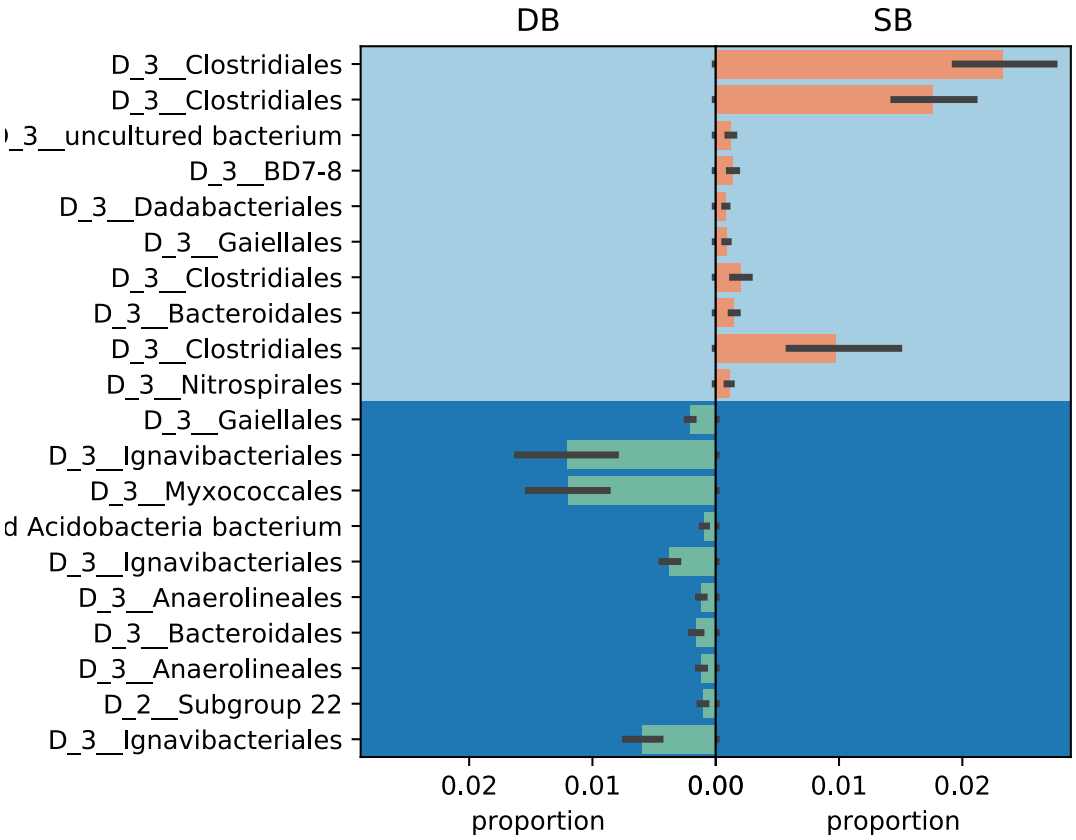


Figure 5.13. Proportion plot of the most differentially abundant ASV taxonomies between Shallow and Deep basin depth cores of the '*y*0' balance.

between sample groups 1 – 3, 1 – 4, 2 – 4, and 3 – 4 (Appendix E.7A). The genus *Clostridiisalibacter* spp. had less of a defined depth-related distribution, decreasing from 7.7% at 0.5 cm to 2.39% at 6 cm (average = 4.73 ± 0.58%), before increasing to 6.27% at 7 cm, and maintaining an average abundance of 6.98 ± 0.45% between 7 – 42 cm. AOV showed a statistical difference in *Clostridiisalibacter* spp. abundance across all sample groups, and Tukey HSD indicated a significant difference between sample group pair 2 and 3 (Appendix E.7B). An uncultured *Anaerolineaceae* ASV was the third most abundant taxon, and had little depth-related abundance distribution, where the average abundance between 0 – 4 cm was 3.52 ± 0.1%, which increased significantly to an average of 4.9 ± 0.15% between 5 – 42 cm. AOV did however show a statistical difference in *Anaerolineaceae* abundance between all sample groups, and Tukey HSD showed a statistical difference between sample group 1 and all three other groups (Appendix E.7C). An uncultured ASV from the δ-*proteobacterial* family *Sandaracinaceae* generally decreased with depth, peaking at 5.45% abundance at 1 cm and was lowest at 42 cm with 1.06% abundance. AOV showed a statistical difference in abundance between all sample groups, and Tukey

HSD showed sample group 4 to have a statistically different abundance of *Sandaracinaceae* with all other sample groups (Appendix E.7D). An uncultured ASV from the *Acidobacteria* family *Thermoanaerobaculaceae* Subgroup 23 was most abundant in the middle region of the Shallow Basin depth core, increasing from 1.58% at 0.5 cm to 3.67% at 6 cm, before declining again to 1.38% at 30 cm. AOV showed a statistical difference in abundance across all sample groups, and Tukey HSD showed a statistical difference between sample groups pairs 1 – 2, and 2 – 4 (Appendix E.7E). An uncultured ASV from the δ-*proteobacteria* family *Desulfobulbaceae* increased from 3.8% at 0.5 cm to 5.25% at 1.5 cm before decreasing to 0.35% at 42 cm. Between 0 – 3 cm, the average abundance was 4.8 ± 0.26%, decreasing to 2.57 ± 0.31% between 4 – 14 cm, and decreasing again to an average abundance of 0.68 ± 0.1% between 16 – 42 cm. AOV showed a statistical difference in *Desulfobulbaceae* abundance across all sample groups, and Tukey HSD showed statistical differences between all sample group pairs, excluding between groups 3 and 4 (Appendix E.7F). The *Clostridiaceae* genus *Alkaliphilus* spp. generally increased in abundance with depth, albeit with a sporadic distribution pattern. Between 0 – 3 cm the average abundance was 0.32 ± 0.06%, which increased to 0.79% at 4 cm, and again to an average of 2.29 ± 0.41% between 4 – 10 cm, then decreased to an average of 1.39 ± 0.35% between 12 – 16 cm. At 18 cm, this increased to 5.47%, before decreasing to an average abundance of 2.35 ± 0.49% between 20 – 42 cm. However, AOV showed only a borderline-significant statistical difference in *Alkaliphilus* spp. abundance across all sample groups, with Tukey showing the borderline-significant group pair to be 1 – 3 (Appendix E.7G). The uncultured *Bacteroidetes* BD2-2 ASV appeared to be most abundant in the middle section of the core. The average abundance between 0 – 2 cm was 1.6 ± 0.1%, which increased to an average of 2.14 ± 0.07% between 3 – 20 cm, and decreased to an average of 1.16 ± 0.23% between 25 – 42 cm. AOV showed a statistical difference in *Bacteroidetes* BD2-2 abundance between all sample groups, and Tukey HSD showed differences between sample group pairs 1 – 3, 3 – 4, and 3 – 4 (Appendix E.7H). An ASV from the γ-*proteobacteria* order B2M28 had a relatively consistent abundance between 0 – 25 cm, with an average abundance of 1.9 ± 0.06%, and this dropped to an average abundance of 0.9 ± 0.15% between 30 – 42 cm. AOV showed a statistical difference in γ-proteobacteria B2M28 abundance across all sample groups, and Tukey HSD showed a statistical difference between group pairs 2 – 4, and 3 – 4 (Appendix E.7I). The *Clostridia* genus *Fusibacter* spp.

was confined to the upper-most region of the core, where the abundance ranged from 2.66% at 1 cm to 8.59% at 4 cm (average 5.59 ± 1%), but which then decreased to 1.27% at 5 cm and 0.86% at 6 cm and decreased again to 0.26 ± 0.1% between 7 – 42 cm, including several depth-increments of 0% abundance. AOV showed borderline-significant difference in *Fusibacter* spp. abundance across all sample depth groups, and Tukey HSD showed the borderline group pair to be 1 – 4 (Appendix E.7J).

Six additional ASVs collapsed at genus level that had relatively high abundance (*i.e.*, > 1%) and had specific catabolic function known or associated to that classification are presented in order of descending total abundance. The α-*proteobacteria* genus *Filomicrobium* spp. increased gradually from 0.48% at 3 cm to a peak of 1.97% at 35 cm (Figure 5.15A). AOV showed a statistical difference in *Filomicrobium* spp. across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.7K). An uncultured δ-*proteobacteria* NB1-j ASV increased from 0.71% at 0.5 cm to 1.81% at 9 cm and decreased to 0.53% at 42 cm (Figure 5.15B). AOV showed a statistical difference in δ-*proteobacteria* NB1-j abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2, 2 – 4, and 3 – 4 (Appendix E.7L). An uncultured archaeal *Bathyarchaeia* increased from 0.1% at 1 cm to 1.88% at 5 cm, decreased to 0.78% at 20 cm, and increased to 1.21% at 42 cm (Figure 5.15C). AOV showed a statistical difference in *Bathyarchaeia* abundance across all sample groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix E.7M). An uncultured *Desulfobacteraceae* Sva0081 sediment group ASV increased from 0.97% at 0.5 cm to 1.75% at 4 cm, decreased generally to 0.94% at 18 cm, and decreased to 0.55% at 42 cm (Figure 5.15D). AOV showed a statistical difference in *Desulfobacteraceae* Sva0081 abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 2 – 3 and 2 – 4 (Appendix E.7N). The *Fusobacteria* genus *Psychrilyobacter* spp. increased from 0.14% at 0.5 cm to 2.96% at 6 cm, decreased to 0.8% at 10 cm, increased to 3.53% at 12 cm, decreased to 0.11% at 18 cm, increased to 2.25% at 20 cm, and decreased to 0.13% at 35 cm (Figure 5.15E). AOV showed a statistical difference in *Psychrilyobacter* spp. abundance across all sample depth groups and Tukey HSD showed a statistical difference between group pair 3 – 4 (Appendix E.7O). The γ-*proteobacteria* genus *Thiogranum* spp. increased

progressively from 0.26% at 0.5 cm to 1.79% at 35 cm (Figure 5.15F). AOV showed a statistical difference in *Thiogranum* spp. across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.7P).

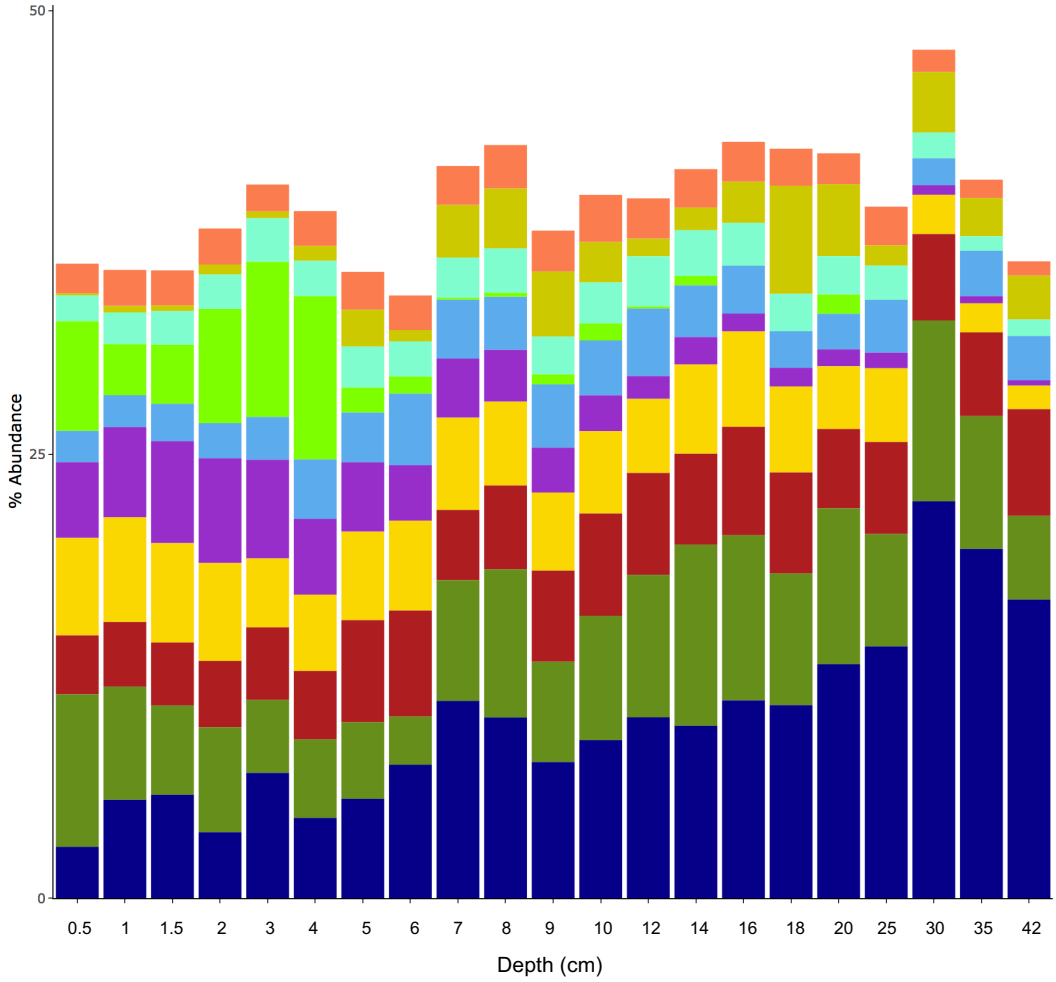


Figure 5.14. Shallow Basin: 100% stacked bar chart ASV taxonomy (collapsed to genus level) abundance, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = *Tepidibacter* spp.; dark green = *Clostridiisalibacter* spp.; red = uncultured *Anaerolineaceae*; yellow = uncultured *Sandaracinaceae*; purple = uncultured *Desulfobulbaceae*; blue = *Thermoanaerobaculaceae* Subgroup 23; light green = *Fusibacter* spp.; light blue = *Bacteroidetes* BD2-2; gold = *Alkaliphilus* spp.; orange = uncultured *γ-proteobacteria* B2M28

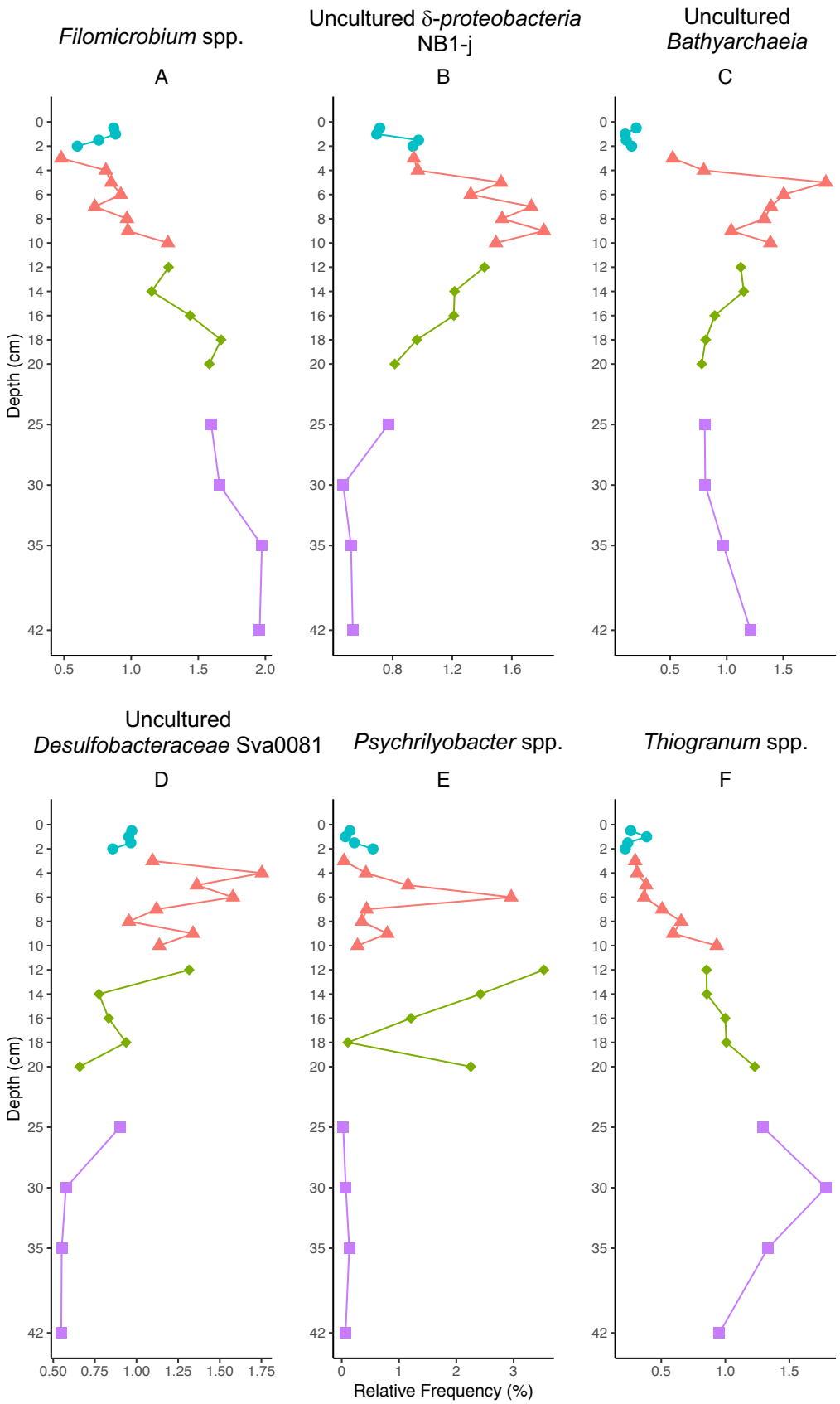


Figure 5.15. Shallow Basin Depth Core: Depth profiles of six additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). *Filomicrobium* spp.; (B). uncultured δ-*proteobacteria* NB1-j; (C). uncultured *Bathyarchaeia*; (D). Uncultured *Desulfobacteraceae* Sva0081; (E). *Psychrilyobacter* spp.; (F). *Thiogranum* spp.

5.3.5.2 *Deep Basin Depth Core Taxonomy*

Firstly, it was observed that the split between bacterial and archaeal abundances in the sample depths ranged from 99.66% bacterial and 0.28% archaeal at 1.5 cm to 80.88% bacterial and 17.12% archaeal (remainder unassigned) at 5 cm. The ten most abundant taxonomic groups when ASVs were collapsed to genus level are presented in Figure 5.16, and these ten taxa accounted for between 23% (0.5 cm) and 37% (18 cm) of the entire community assemblages.

The relative abundance of the ten most abundant ASVs in the deep basin showed a more progressive depth-related pattern (Figure 5.16) with either progressive increases or decreases with depth or discernible subsurface peaks at specific depths, and abundances are presented below in descending order of total abundance.

The most abundant collapsed ASV taxonomic group was related to the *Anaerolineaceae*. Between 0 – 1.5 cm, the average abundance was 1.94 ± 0.07%, which increased to 3.35% at 2 cm and 4% at 4 cm. Between 5 – 47 cm, the average abundance was 5.16 ± 0.14%. AOV showed a statistical difference in *Anaerolineaceae* abundance between all groups, and Tukey HSD showed a statistical difference between sample group 1 with the three other groups (Appendix E.8A). An uncultured ASV from the δ-*proteobacteria* family *Sandaracinaceae* decreased in abundance consistently with depth, with an abundance of 8.08% at 0.5 cm progressing incrementally to 0.17% at 47 cm. Subsequently, AOV and Tukey HSD showed statistical differences in abundance across all sample groups and pairs (Appendix E.8B). An ASV from the *Bacteroidetes* BD2-2 family increased in abundance between 0 – 5 cm, from 2.13% at 0.5 cm to 5.63% at 5 cm. This group then also decreased consistently with depth, from 4% at 6 cm to 0.5% at 47 cm (with peaks of 4.91% at 8 cm and 3.47% at 14 cm). AOV resulted in a statistical difference in *Bacteroidetes* BD2-2 abundance across all sample groups, while Tukey HSD showed statistical differences between all sample group pairs, with the exception of pair 2 and 3 (Appendix E.8C). An ASV from the *Thermoanaerobaculaceae* Subgroup 23 increased in abundance between 0 – 6 cm, from 1.73% to 3.86%. Abundance then decreased with depth from 3.77% at 7 cm to 0.55% at 47 cm. AOV showed a statistical difference in *Thermoanaerobaculaceae* Subgroup 23 abundance between all sample groups, and Tukey HSD showed statistical difference between sample

group pairs 1 – 4, 2 – 3, and 2 – 4 (Appendix E.8D). An uncultured *Clostridiaceae* had an average abundance of 0.06 ± 0.01% between 0 – 2 cm, which increased to 0.45 ± 0.11% between 3 – 5 cm and increased again from 1.27% at 6 cm to 4.59% at 18 cm (average = 2.96 ± 0.37%). Between 20 – 47 cm, the average abundance was 3.94 ± 0.13%. AOV showed a statistical difference across all sample groups, and Tukey HSD showed a statistical difference between all sample group pairs, with the exception of pair 3 and 4 (Appendix E.8E). An archaeal ASV from the *Bathyarchaeia* had a low relative abundance initially which increased with depth, with an average of 0.08 ± 0.0% between 0 – 1.5 cm, 0.48 ± 0.09% between 2 – 5 cm, and increased to 0.72% at 6 cm to 6.74% at 47 cm. AOV showed a statistical difference in *Bathyarchaeia* abundance and Tukey HSD showed a statistical difference between all sample group pairs excluding pair 1 and 2 (Appendix E.8F). An uncultured ASV from the δ-*proteobacterial* family *Syntrophobacteraceae* was generally most abundant in the middle region of the core and was characterised by numerous specific depth-increments of increased abundance. Between 0 – 5 cm, the average abundance was 1.07 ± 0.16%, which increased to 4 ± 0.53% between 6 – 9 cm, decreased to 2.49 ± 0.09% between 10 – 16 cm, increased to 7.53% at 18 cm, and then decreased from 2.95% at 20 cm to 0.22% at 47 cm (with an increase to 5.08% at 27 cm). AOV showed borderline-significant difference in *Syntrophobacteraceae* abundance across all sample groups, with Tukey showing the borderline-significant difference to be between group pair 1 – 3 (Appendix E.8G). An uncultured γ-*proteobacteria* was almost exclusively confined to the upper half of the core, where it increased in abundance from 4.13% at 0.5 cm to 7.09% at 1.5 cm, decreased to 0.61% at 8 cm (average 2.88 ± 0.62%), and had an average abundance of 0.23 ± 0.04% between 9 – 47 cm. AOV showed a statistical difference in the γ-*proteobacteria* ASV abundance and Tukey HSD showed a statistical difference across all sample group pairs excluding pair 3 and 4 (Appendix E.8H). An uncultured *Ignavibacteriales* PHOS-HE36 ASV increased consistently with depth from 0.49% at 0.5 cm to 3.56% at 18 cm (average 2 ± 0.28%) and decreased to 0.76% at 47 cm (average 1.83 ± .022%). AOV showed a statistical difference in abundance, and Tukey HSD showed statistical differences in abundance between sample group pairs 1 – 2, 1 – 3, and 3 – 4 (Appendix E.8I). The *Tepidibacter* ASV increased from 0.32% and 0.5 cm to 5.08% at 16 cm (average 1.9 ± 0.43%) and decreased to 0.88% at 47 cm (average 1.74 ± 0.29%). AOV showed a statistical difference in *Tepidibacter*

abundance and Tukey HSD showed a statistical difference in abundance between sample group pairs 1 – 2, 1 – 3, 2 – 3, and 3 – 4 (Appendix E.8J).

Six additional ASVs collapsed at genus level that had relatively high abundance (*i.e.*, > 1%) and had specific catabolic function known or associated to that classification are presented in order of descending total abundance. An uncultured *Desulfobulbaceae* was most abundant in the upper region of the core, with an average abundance of 3.66 ± 0.22% between 0 – 4 cm, which then decreased to 0.71% at 14 cm (average 1.56 ± 0.34%) and decreased again to 0.62 ± 0.15% between 16 – 47 cm (Figure 5.17A). AOV showed a statistical difference in *Desulfobulbaceae* abundance and Tukey HSD showed a statistical difference in abundance between sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.8K). An uncultured δ-*proteobacteria* NB1-j ASV generally increased with depth, increasing from 1.22% at 0.5 cm to 2.44% at 10 cm, which then decreased to an average abundance of 2.18 ± 0.08% between 12 – 32 cm, decreasing again to an average abundance of 0.94 ± 0.1% between 37 – 47 cm (Figure 5.17B). AOV showed a statistical difference in δ-*proteobacteria* NB1-j abundance across sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, and 3 – 4 (Appendix E.8L). The *Desulfatiglans* spp. increased gradually from 0% at 0.5 cm to 1.47% at 12 cm, decreased to 1.1% at 14 cm, increased to 2.19% at 27 cm, and decreased to 1.62% at 47 cm (Figure 5.17C). AOV showed a statistical difference in *Desulfatiglans* spp. abundance across all sample depth groups and Tukey HSD showed statistical differences between all pairs of groups excluding 3 – 4 (Appendix E.8M). The *Ignavibacterium* spp. followed a very similar pattern to the *Desulfatiglans* spp., increasing from 0% at 0.5 cm to 1.79% at 14 cm, decreased to 0.53% at 27 cm, and increased to 2.57% at 42 cm (Figure 5.17D). AOV showed a statistical difference in *Ignavibacterium* spp. abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, and 2 – 4

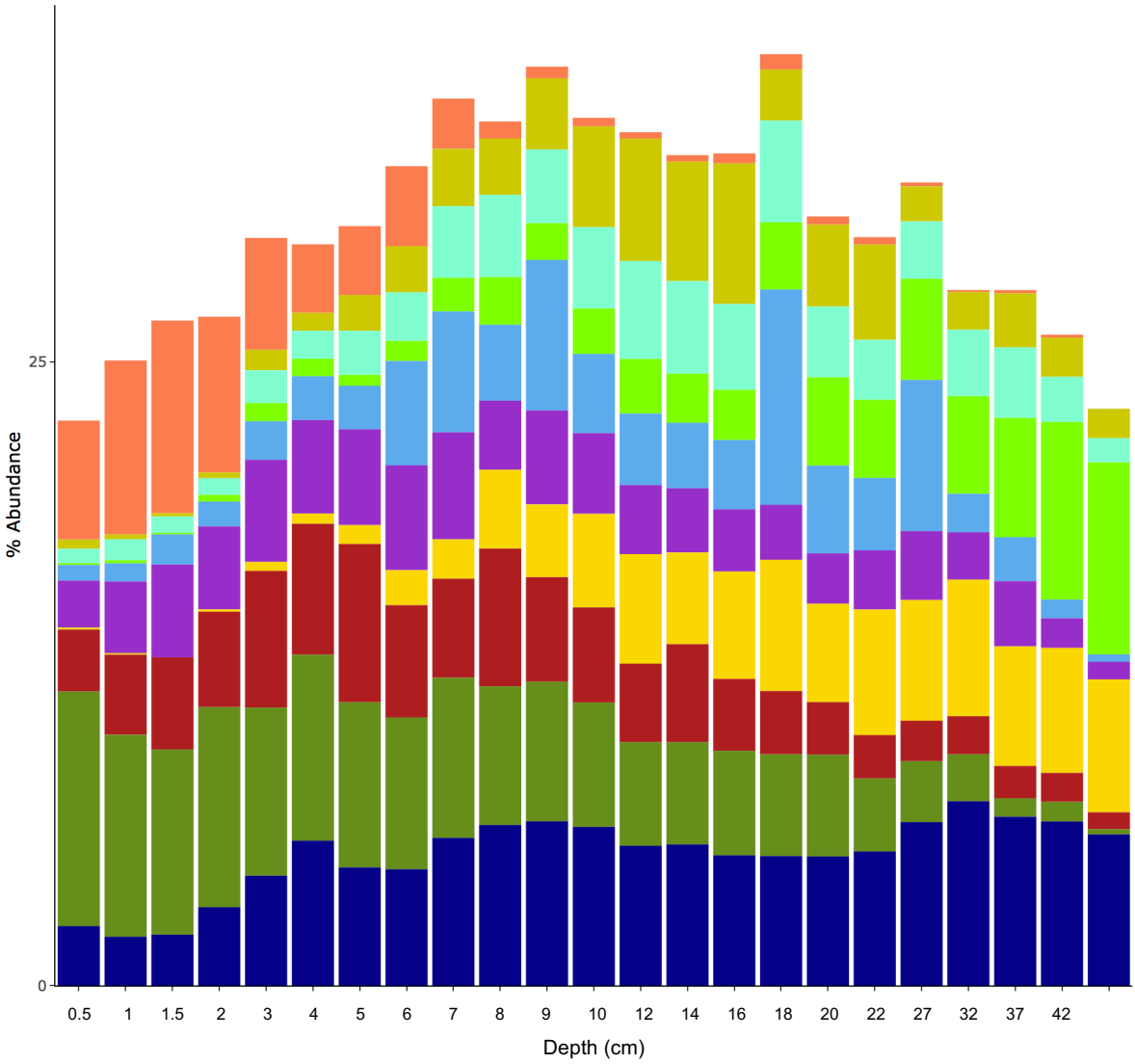


Figure 5.16. Deep Basin: 100% stacked bar chart ASV taxonomy (collapsed to genus level) abundance, showing the ten most abundant ASV genera. Size of bar reflects normalised abundance of that genera as a proportion of the total ASV genera assemblage within each sample. Dark blue = uncultured *Anaerolineaceae*; dark green = uncultured *Sandaracinaceae*; red = *Bacteroidetes* BD2-2; yellow = uncultured *Clostridiaceae*; purple = *Thermoanaerobaculaceae* Subgroup 23; blue = uncultured *Syntrophobacteraceae*; green = uncultured *Bathyarchaeia*; light blue = *Ignavibacteriales* PHOS-HE36; gold = *Tepidibacter* spp.; orange = uncultured γ-*proteobacteria*.

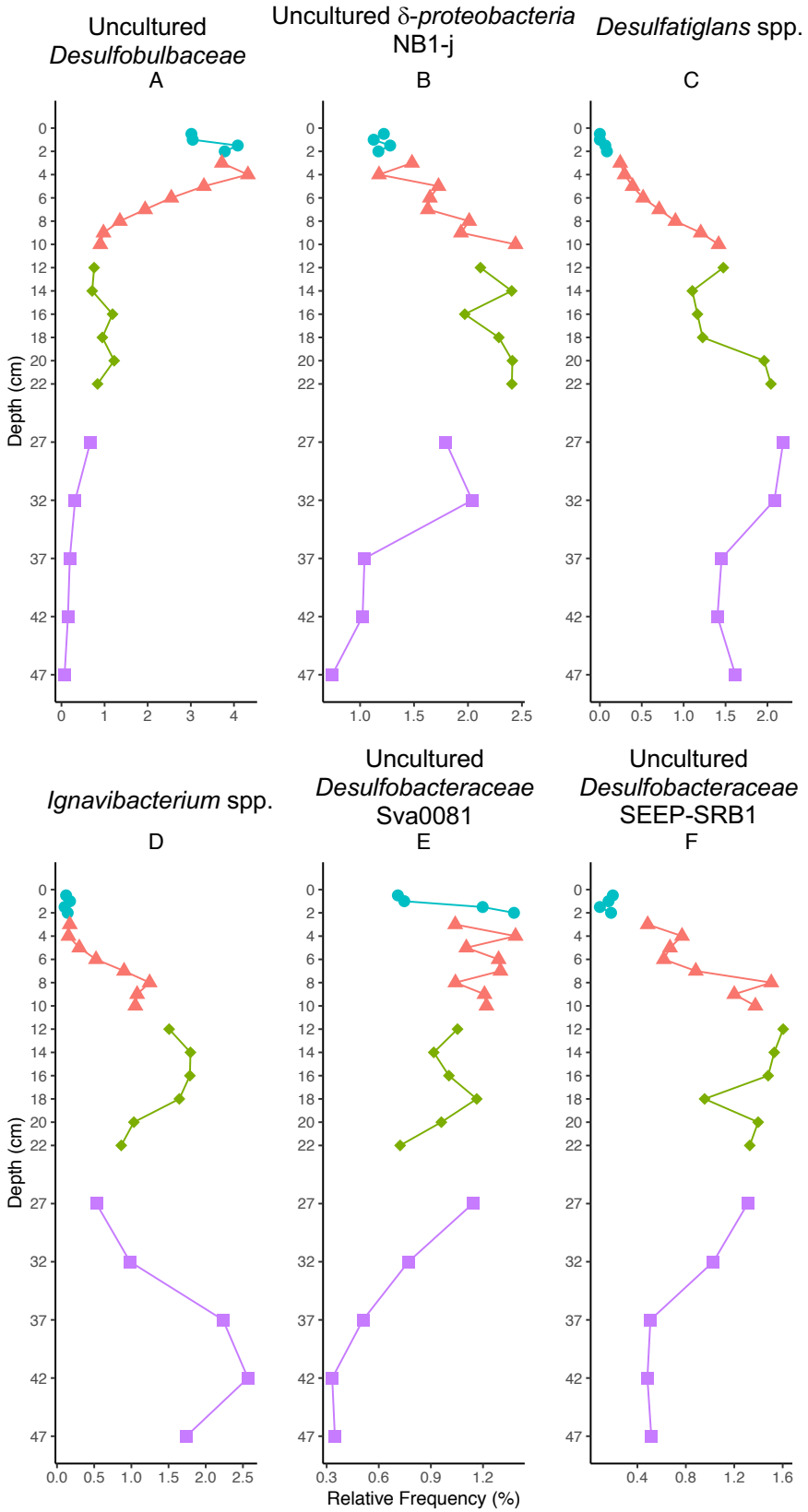


Figure 5.17. Deep Basin Depth Core: Depth profiles of three additional microbial taxonomic groups of interest. Values represent relative frequency of the ASV taxonomy collapsed at genus level for each depth sample. Colours and points denote sample depth group. (A). uncultured *Desulfobulbaceae*; (B). Uncultured δ-*proteobacteria* NB1-j.; (C). *Desulfatiglans* spp.; (D). *Ignavibacterium* spp.; (E). Uncultured *Desulfobacteraceae* Sva0081; (F). Uncultured *Desulfobacteraceae* SEEP-SRB1.

(Appendix E.8N). An uncultured *Desulfobacteraceae* Sva0081 increased from 0.71% at 0.5 cm to 1.38% at 2 cm, and generally decreased with depth to 0.35% at 47 cm (Figure 5.17E). AOV showed a statistical difference in *Desulfobacteraceae* Sva0081 abundance across all sample depth groups and Tukey HSD showed a statistical difference between group pair 2 – 4 (Appendix E.8O). An uncultured *Desulfobacteraceae* SEEP-SRB1 increased from 0.2% at 0.5 cm to 1.51% at 8 cm, decreased gradually to1.03% at 32 cm, and decreased to 0.48% at 42 cm (Figure 5.17F). AOV showed a statistical difference in *Desulfobacteraceae* SEEP-SRB1 abundance across all sample depth groups and Tukey HSD showed statistical differences between group 1 and all other groups and between group pair 3 – 4 (Appendix E.8P).

5.3.6 Tax4Fun Results

A total of 6,572 genes were predicted by Tax4Fun in the Shallow Basin core, and 6,569 genes in the Deep Basin core. Target gene abundances reflect the percentage normalised abundance of the gene at a given sample depth.

5.3.6.1 *Shallow Basin Depth Core Inferred Functional Gene Abundance*

The five functional genes targeted during 'end point' PCR were identified in the output functional gene abundance profile inferred by Tax4Fun from the Shallow Basin depth core. The ammonia monooxygenase *amoA* (not distinguished between archaeal and bacterial variants of the gene) gene was inferred to have a low abundance throughout the core, with an average abundance of 0.07 ± 0.01% between 0 – 42 cm, with a peak abundance of 0.17% at 1.5 cm (Figure 5.18A). AOV showed a statistical difference in *amoA* abundance between all sample groups, and Tukey HSD showed a statistical difference between sample group pairs 1 – 2, 1 – 3, and 2 – 4 (Appendix E.9A). The denitrifying nitrite reductase *nirK* gene initially increased from 0.7% at 0.5 cm to 0.1% at 1.5 cm, and decreased to 0.05% at 3 cm, and then the abundance steadily increased from 0.06% at 4 cm to 0.09% at 42 cm (average 0.7 ± 0.03%) (Figure 5.18B). AOV showed a statistical difference in *nirK* abundance and Tukey HSD showed statistical differences in pairs 1 – 2, 1 – 3, and 2 – 4 (Appendix E.9B). The other denitrifying nitrite reductase *nirS* increased from 1.1% at 0.5 cm to 2.9% at 16 cm (average 1.7 ± 0.1%) and decreased to an average abundance of 2.1 ± 0.1% between 18 – 42 cm (Figure 5.18C). AOV showed a statistical difference between all sample groups and Tukey HSD showed statistical

differences between sample group pairs 1 and all three other groups, and 2 – 3 (Appendix E.9C). The dissimilatory sulfite reductase *dsrA* gene initially increased between from 1.7% at 0.5 cm to 3% at 4 cm, generally decreased from 2.5% at 5 cm to 1% at 16 cm, and between 18 – 42 cm had an average abundance of 1 ± 0.1% (Figure 5.18D). AOV showed a statistical difference in *dsrA* abundances across all sample groups, and Tukey HSD showed statistical differences in sample group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.9D). The *mcrA* gene was absent in most samples, never exceeding 0.017% in two samples (Figure 5.18E). AOV and Tukey did not show statistical differences in *mcrA* abundance (Appendix E.9E). Of the four genes compared between 'end point' PCR band intensity and Tax4Fun inferred fractional abundance (archaeal and bacterial *amoA*, *nirK*, and *nirS*), no positive correlations were found (Figure 5.19A–D). Furthermore, archaeal *amoA*, *nirK*, and *nirS* all had negative ρ values, indicating negative correlation in values (*i.e.*, abundance) – with *nirK* and *nirS* having a statistical negative correlation (*i.e.*, *p* value < 0.05).

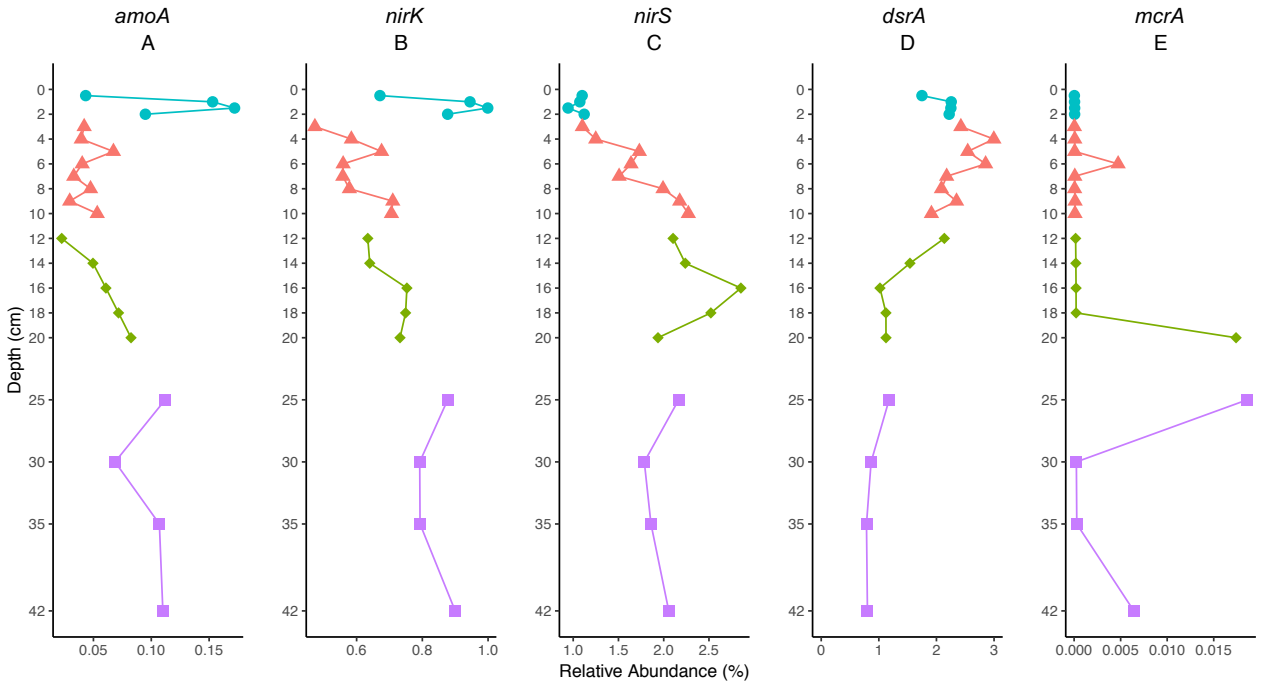


Figure 5.18. Shallow Basin Depth Core: Depth profile of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA*.

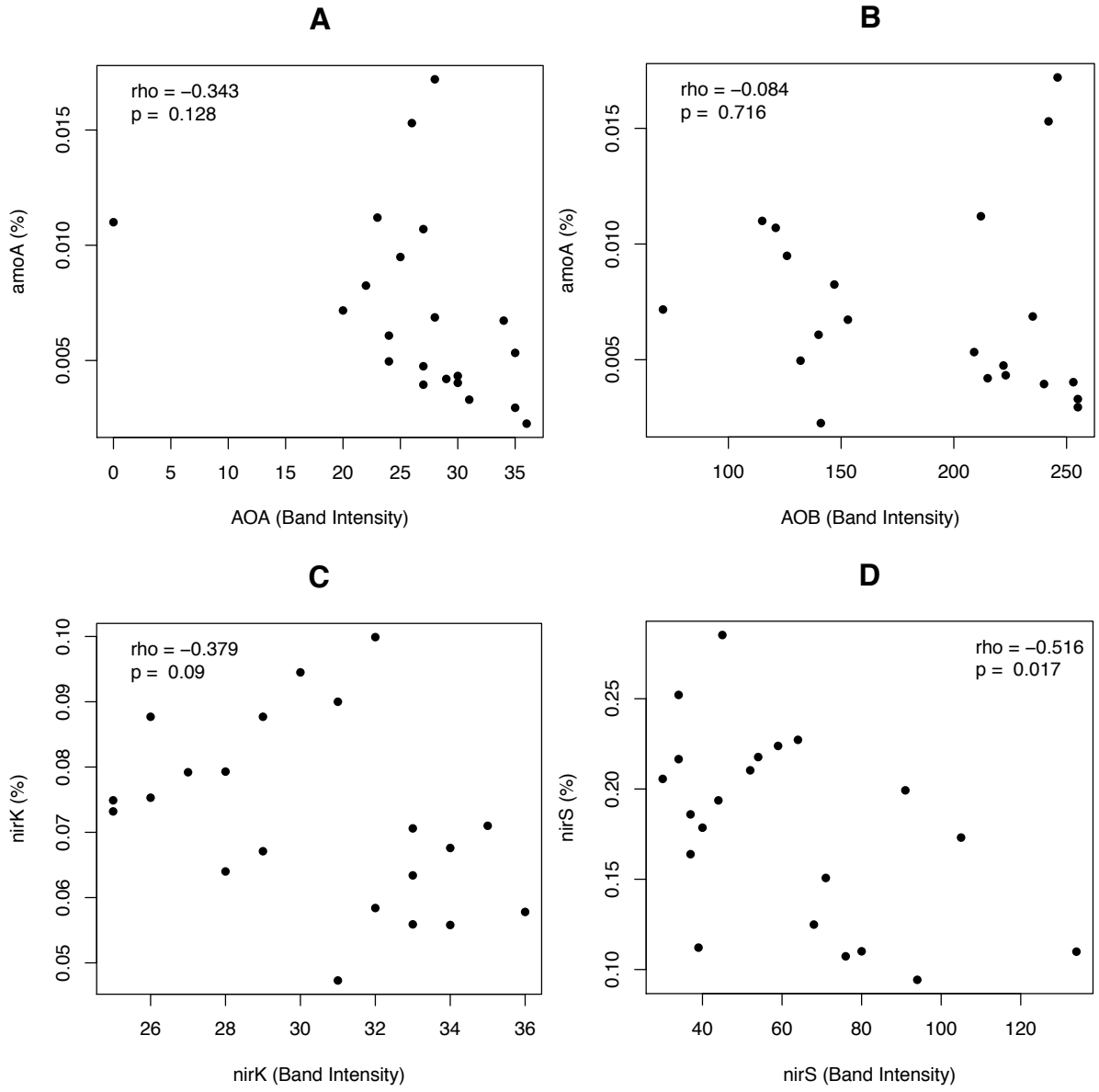


Figure 5.19. Shallow Basin Depth Core: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) (X axis) and the abundance of the same gene as inferred by Tax4Fun (relative abundance) (Y axis). Spearman correlation is represented by ρ ('*rho*') and *p* values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*.

Nitrogen cycling genes were consistently more abundant in the upper region. The *narG*, *napA*, and *norB* all consistently increased between 0.5 and 4 cm (3.5%, 2.1%, and 2% at 0.5 cm to 4.7%, 2.6%, and 2.4% at 4 cm respectively) and decreased gradually to until 20 cm (2.8%, 1.1%, and 1.5% respectively), increased at 25 cm (to 3.5%, 1.4%, and 1.8% respectively) and decreased at 42 cm (2.7%, 0.92%, and 1.5% respectively) (Figures 5.20A, 5.20B, and 5.20C). AOV showed statistical differences in *narG*, *napA*, and *norB* abundances across all sample depth groups, and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 –

4, 2 – 3, and 2 – 4 (Appendix E.10A, 5.11B, and 5.11C). The *nosZ* gene increased from 1.6% at 1 cm to 2% at 2 cm and decreased progressively to 0.9% at 35 cm (Figure 5.20D). AOV showed a statistical difference in *nosZ* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, and 2 – 4 (Appendix E.10D). The DNRA *nrfA* generally decreased from 3.2% at 2 cm to 1.5% at 35 cm (Figure 5.20E). AOV showed a statistical difference in *nrfA* abundance across all sample groups and Tukey HSD showed statistical differences between all sample group pairs except 1 – 2 (Appendix E.10E). The nitrogen-fixating *nifD* increased from 2.3% at 0.5 cm to 3.2% at 5 cm and decreased gradually to 1.4% at 35 cm (Figure 5.20F). AOV showed a statistical difference in *nifD* across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.10F).

Other target genes included a mix of aerobic and anaerobic functional genes and are not presented in a specific order. The Wood-Ljungdahl anaerobic carbon fixating gene *acsB* increased from 2.6% at 0.5 cm to 5% at 4 cm and decreased gradually to 1.5% at 35 cm (Figure 5.21A). AOV showed a statistical difference in *acsB* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 4, 2 – 3, and 2 – 4 (Appendix E.11A). The dissimilatory sulfate reduction pathway gene *aprA* increased from 2.9% at 0.5 cm to 4.7% at 4 cm and decreased gradually to 1.5% at 42 cm (Figure 5.21B). AOV showed a statistical difference in *aprA* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 4, 2 – 3, and 2 – 4 (Appendix E.11B). The methane-oxidising *mmoX* increased decreased from 0.14% at 0.5 cm to 0.11% at 2 cm, increased gradually to 0.23% at 42 cm (Figure 5.21C). AOV showed a statistical difference in *mmoX* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.11C). Another methane-oxidising gene, the *mxaF* followed the same pattern as the *mmoX*, ranging from 0.25% at 3 cm to 0.6% at 35 cm (Figure 5.21D). AOV showed a statistical difference in *mxaF* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.11D). The NiFe hydrogenase *hydB* had a mixed and relatively low distribution with depth, ranging from 0.05% at 1.5 cm to 0.09% at 2 cm (Figure 5.21E). AOV and Tukey did not show statistical differences in *hydB*

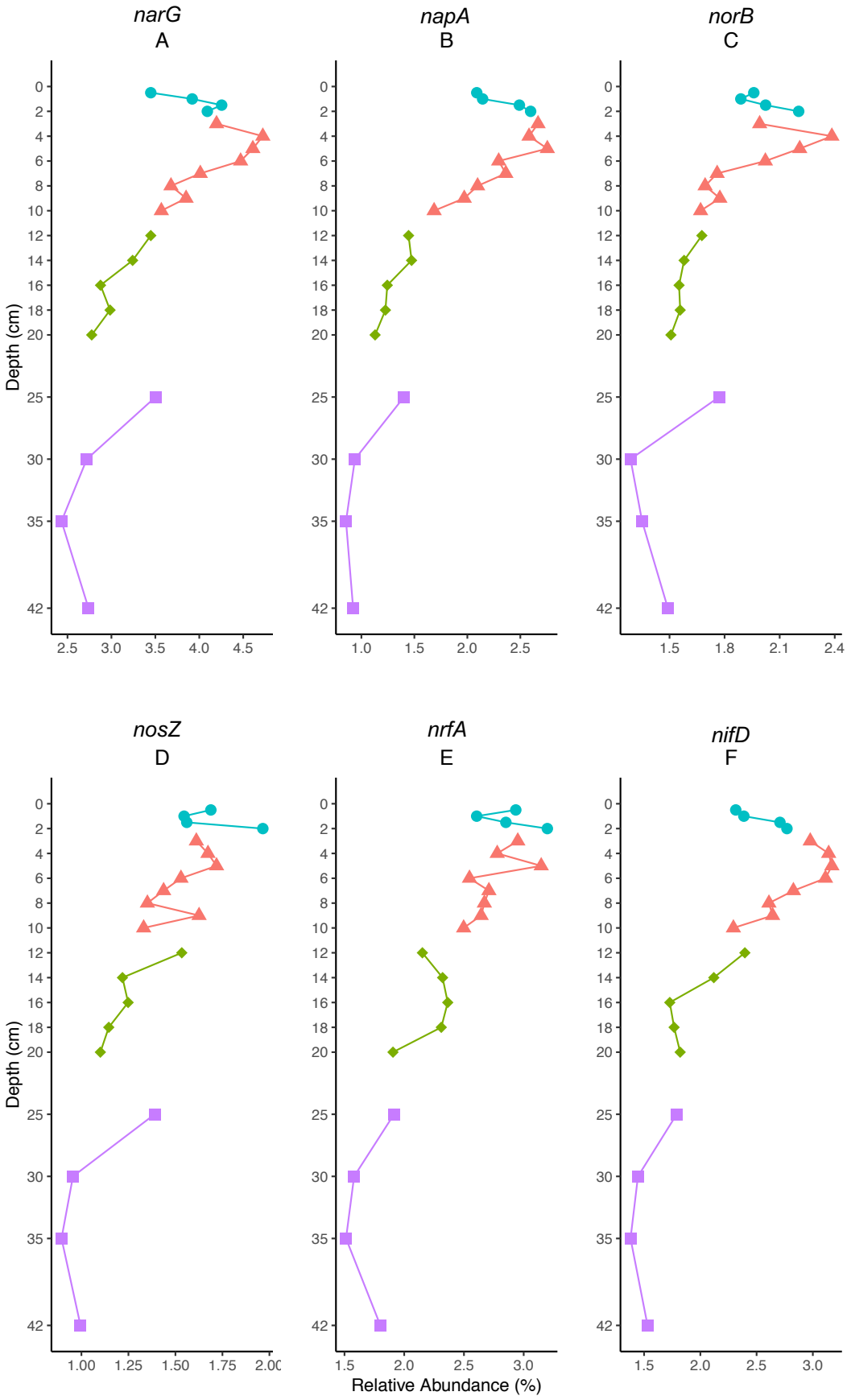


Figure 5.20. Shallow Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD*.

abundances (Appendix E.11E). The methanotrophic gene *mtdB* followed a similar pattern to the *mmoX* and *mxaF*, decreasing from 0.25% at 0.5 cm to 02% at 2 cm, increasing to 0.3% at 42 cm (Figure 5.21F). AOV showed a statistical difference in *mtdB* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 4, 2 – 3, and 2 – 4 (Appendix E.11F). The methanogenic and methane-oxidising *hdrA* increased from 16% at 0.5 cm to 31.2% at 4 cm and decreased to 5.9% at 35 cm (Figure 5.21G). AOV showed a statistical difference in *hdrA* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 4, 2 – 3, and 2 – 4 (Appendix E.11G). Another anaerobic carbon fixating gene *cooS* followed the same pattern as the *acsB* and increased from 4.7% at 0.5 cm to 8% at 4 cm, and decreased to 2.7% at 42 cm (Figure 5.21H). AOV showed a statistical difference in *cooS* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 4, 2 – 3, and 2 – 4 (Appendix E.11H).

5.3.6.2 *Deep Basin Depth Core Inferred Functional Gene Abundance*

The five genes targeted by 'end point' PCR were identified in the output functional gene abundance profile inferred by Tax4Fun. The *amoA* gene generally remained low, with an average abundance of 0.057 ± 0.009% between 0.5 – 42 cm, increasing to 0.26% at 47 cm (Figure 5.22A). AOV and Tukey did not show statistical differences in *amoA* abundance across or between sample groups (Appendix E.12A). The *nirK* decreased from 1% at 0.5 cm to 0.5% at 9 cm and increased to 1.4% at 47 cm (Figure 5.22B). AOV showed a borderline-significant difference in *nirK* abundance across sample groups (Appendix E.12B). The *nirS* gene increased gradually with depth from 0.9% at 0.5 cm to 4% at 47 cm (Figure 5.22C). AOV showed a statistical difference in *nirS* abundance across all sample groups and Tukey HSD showed statistical differences between group 4 and all other groups (Appendix E.12C). The *dsrA* gene increased from 1.8% at 0.5 cm to 5% at 20 cm and decreased to 1.6% at 47 cm (Figure 5.22D). AOV showed a statistical difference in *dsrA* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 2 – 3 (Appendix E.12D). The *mcrA* was nearly absent from all but two samples and never exceeded 0.015% (Figure 5.22E). AOV and Tukey did not show statistical differences in *mcrA* abundance across or between sample groups (Appendix E.12E). The *amoA* gene did not have a statistically significant positive correlation for either archaeal or bacterial sources (Figures 5.23A and 5.23B). The

nirK gene had a statistically significant positive correlation between 'end point' PCR band intensity and Tax4Fun inferred fractional gene abundance, albeit not a strong correlation (Figure 5.23C). Similar to the Shallow Basin, the *nirS* had a statistically significant negative correlation between band intensity and Tax4Fun abundance (Figure 5.23D).

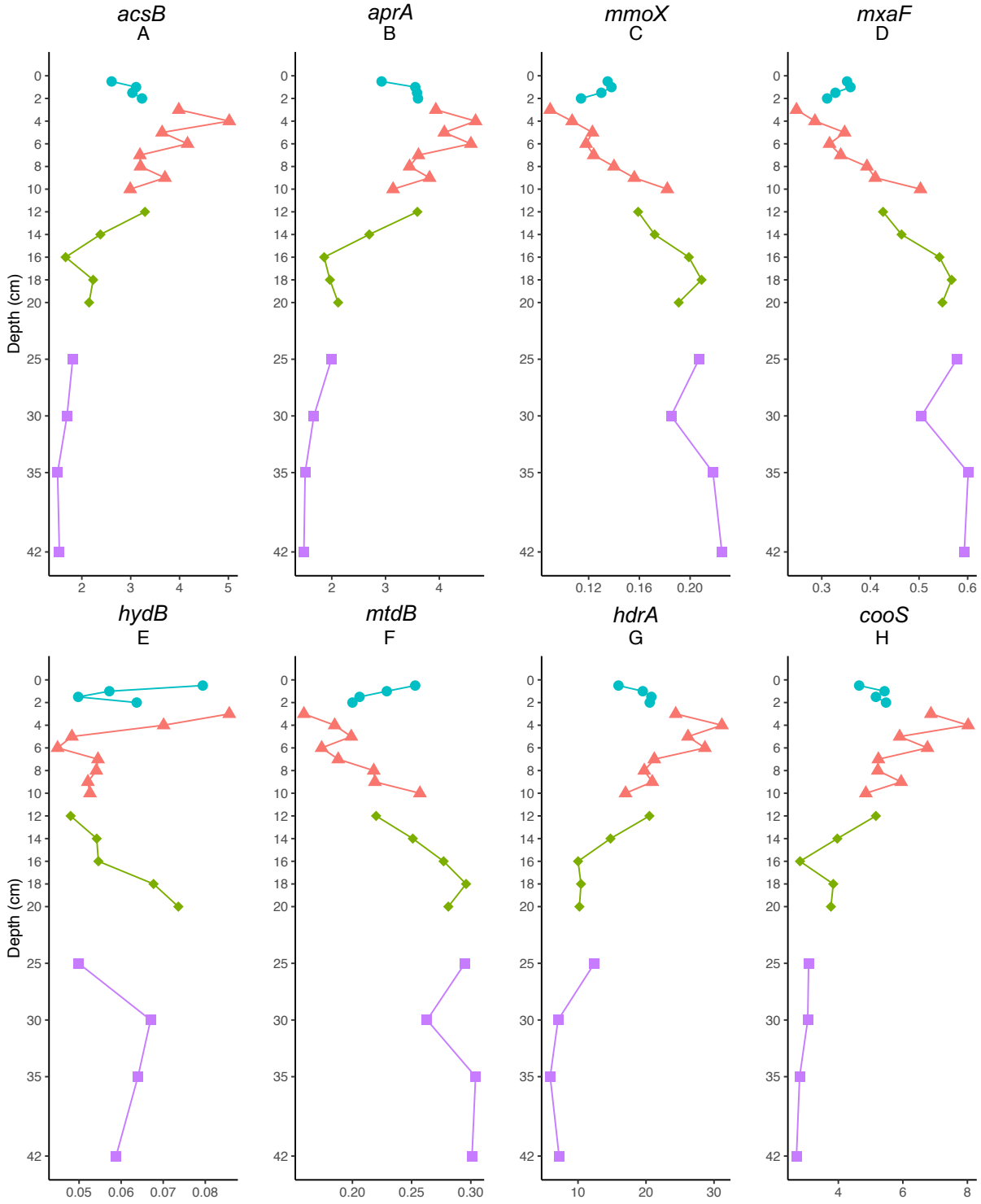


Figure 5.21. Shallow Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS*.

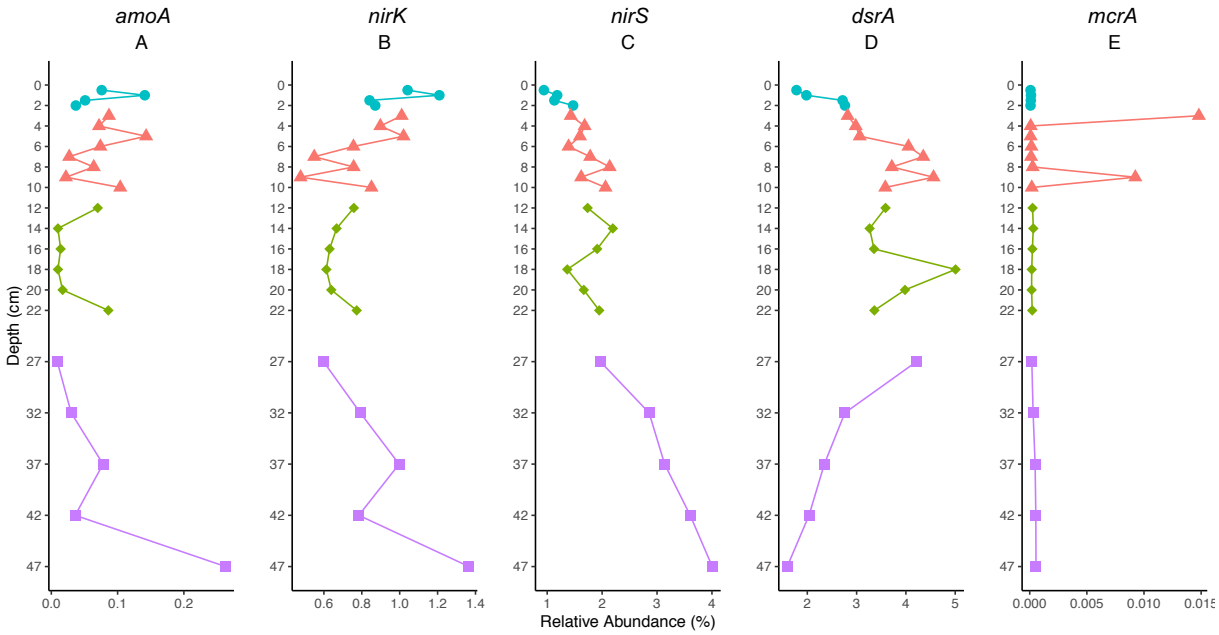


Figure 5.22. Deep Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of genes targeted by 'end point' PCR, as a function of depth in cm. Colour shapes refer to sample depth group. (A). *amoA*; (B). *nirK*; (C). *nirS*; (D). *dsrA*; (E). *mcrA*.

The nitrogen cycling genes appeared to be most abundant in the middle region of the Deep Basin depth core. Similar to the Shallow Basin, the *narG*, *napA*, and *norB* all had very similar patterns of distribution, increasing between 0.5 and 9 cm, decreasing at 10, peaking at 20 cm, and decreasing to 47 cm, ranging from 3.1%, 1.7%, and 2.3% at 0.5 cm to 7.9%, 3.9%, and 5.3% at 18 cm respectively (Figures 5.24A, 5.24B, and 5.24C). AOV showed a statistical difference in *narG* abundance across all sample groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix E.13A).

AOV showed a statistical difference in *napA* abundance across all sample groups and Tukey HSD showed a statistical difference between group pair 2 – 4 (Appendix E.13B). AOV showed a borderline-significant difference in *norB* abundance across all sample groups, and Tukey HSD showed the borderline-significant difference to be between group pair 1 – 3 (Appendix E.13C). The *nosZ* had an average abundance of 2.5 ± 0.06% between 0.5 – 16 cm, increased to 3.6% at 18 cm, and decreased to 0.9% at 47 cm (Figure 5.24D). AOV showed a statistical difference in *nosZ*

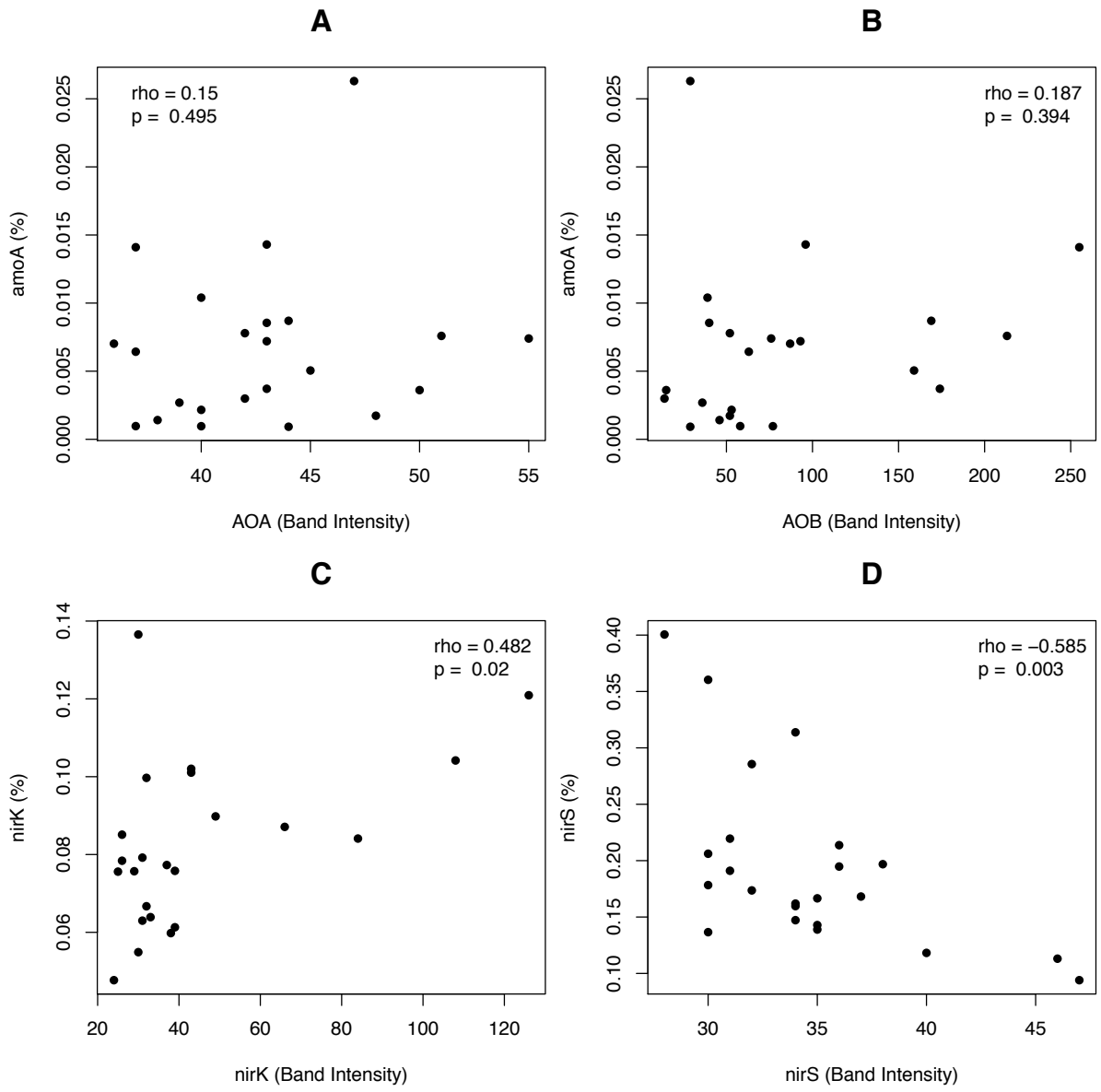


Figure 5.23. Deep Basin: XY scatterplots comparing abundances of the four 'end point' PCR functional gene targets as determined by BioNumerics (band intensity) (X axis) and the abundance of the same gene as inferred by Tax4Fun (relative abundance) (Y axis). Spearman correlation is represented by ρ ('*rho*') and *p* values. (A). Archaeal *amoA*; (B). Bacterial *amoA*; (C). *nirK*; (D). *nirS*.

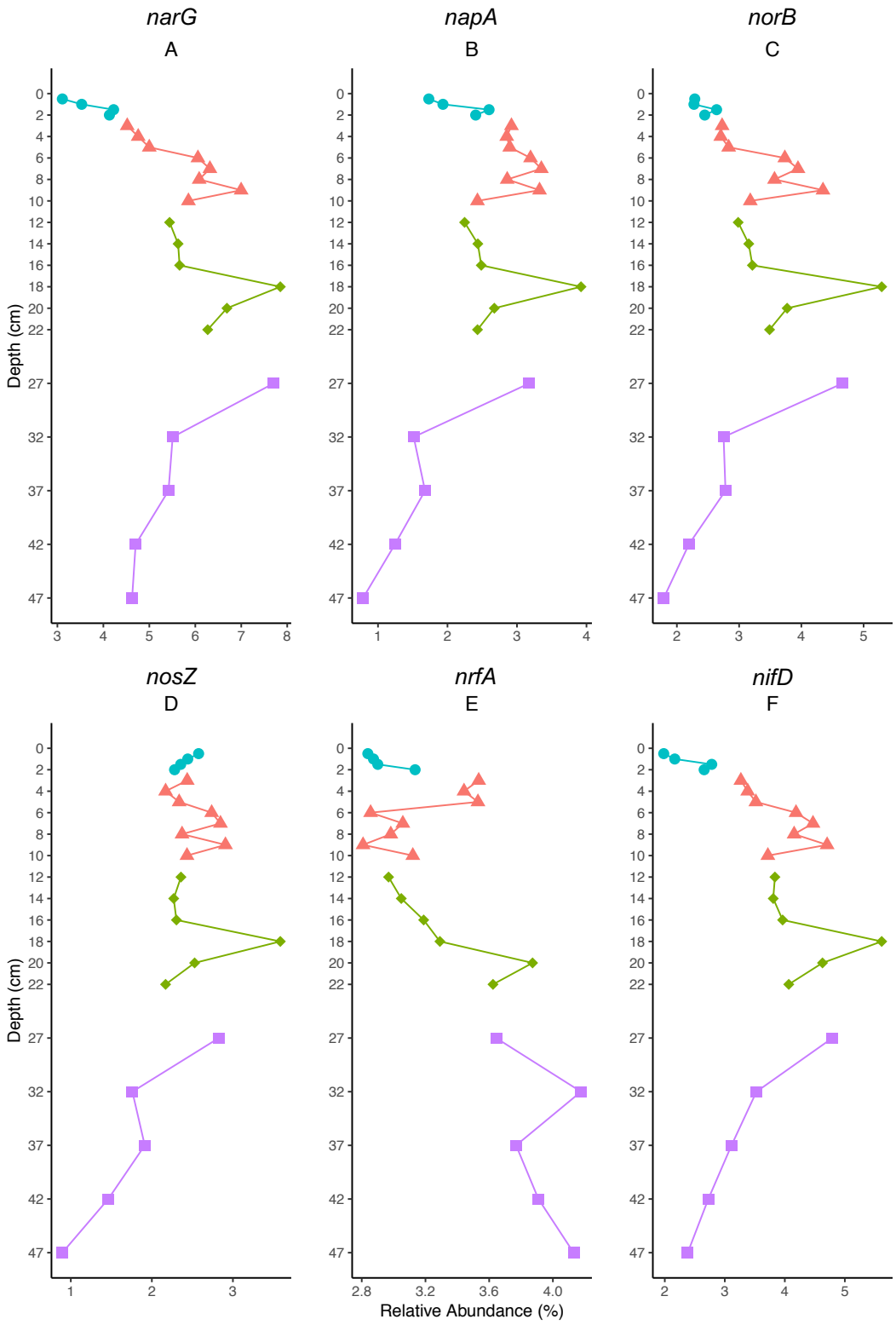


Figure 5.24. Deep Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of nitrogen metabolism genes as a function of depth in cm. Colour shapes refer to sample depth group. (A). *narG*; (B). *napA*; (C). *norB*; (D). *nosZ*; (E). *nrfA*; (F). *nifD*.

abundance across all sample groups and Tukey HSD showed a statistical difference between group pair 2 – 4 (Appendix E.13D). The *nrfA* increased from 2.8% at 0.5 cm to 3.5% at 5 cm, decreased to 2.9% at 6 cm, and increased to 4.1% at 47 cm (Figure

5.24E). AOV showed a statistical difference in *nrfA* abundance across all sample groups and Tukey HSD showed statistical differences between group 4 and all other groups (Appendix E.13E). The *nifD* followed a similar pattern to the *narG*, *napA*, and *norB* genes increasing from 2% at 0.5 cm to 4.7% at 9 cm, decreasing to 3.7% at 10 cm, peaking at 5.6% at 18 cm, and decreasing to 2.4% at 47 cm (Figure 5.24F). AOV showed a statistical difference in *nifD* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 1 – 3 (Appendix E.13F).

The *acsB*, *aprA*, *hdrA*, and *cooS* genes all followed the same distribution patterns as the *narG*, *napA*, *norB*, and *nifD* genes. The *acsB* was 2.1% at 0.5 cm, 7.8% at 18 cm, and 2.6% at 47 cm (Figure 5.25A). AOV showed a statistical difference in *acsB* abundance across all sample groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix E.14A). The *aprA* gene was 2.8% at 0.5 cm, 7.5% at 18 cm, and 2.5% at 47 cm (Figure 5.25B). AOV showed a statistical difference in *aprA* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 1 – 3 (Appendix E.14B). The *hdrA* gene was 15.6% at 0.5 cm, 79.6% at 18 cm, and 25.3% at 47 cm (Figure 5.25G). AOV showed a statistical difference in *hdrA* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 1 – 3 (Appendix E.14G). The *cooS* gene was 3.1% at 0.5 cm, 9% at 18 cm, and 4.4% at 47 cm (Figure 5.25H). AOV showed a statistical difference in *cooS* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 1 – 3 (Appendix E.14H). The *hydB* did not appear to have a depth-related distribution of abundance and was low throughout ranging from 0.01% at 18 cm to 0.03% at 42 cm (Figure 5.25E). AOV and Tukey did not show statistical differences in the *hydB* abundance across or between sample groups (Appendix E.14E). The *mmoX*, *mxaF*, and *mtdB* all had similar patterns that saw them generally decrease from 0.5 to 18 cm, and generally increase to 47 cm. The *mmoX* was 0.16% at 0.5 cm, 0.07% at 18 cm and 0.16% at 47 cm (Figure 5.25C). AOV showed a statistical difference in *mmoX* abundance across all sample groups and Tukey HSD showed statistical differences between group pairs 1 – 2 and 1 – 3 (Appendix E.14C). The *mxaF* was 0.39% at 0.5 cm, 0.17% at 18 cm, and 0.37% at 47 cm (Figure 5.25D). AOV showed a statistical difference in *mxaF* abundance across all sample groups and Tukey HSD showed statistical differences

between group pairs 1 – 2 and 1 – 3 (Appendix E.14D). The *mtdB* was 0.27% at 0.5 cm, 0.08 at 18 cm, and 0.17% at 47 cm (Figure 5.25F). AOV showed a statistical difference in *mtdB* abundance across all sample groups and Tukey HSD showed statistical differences between group 1 and all other groups (Appendix E.14C).

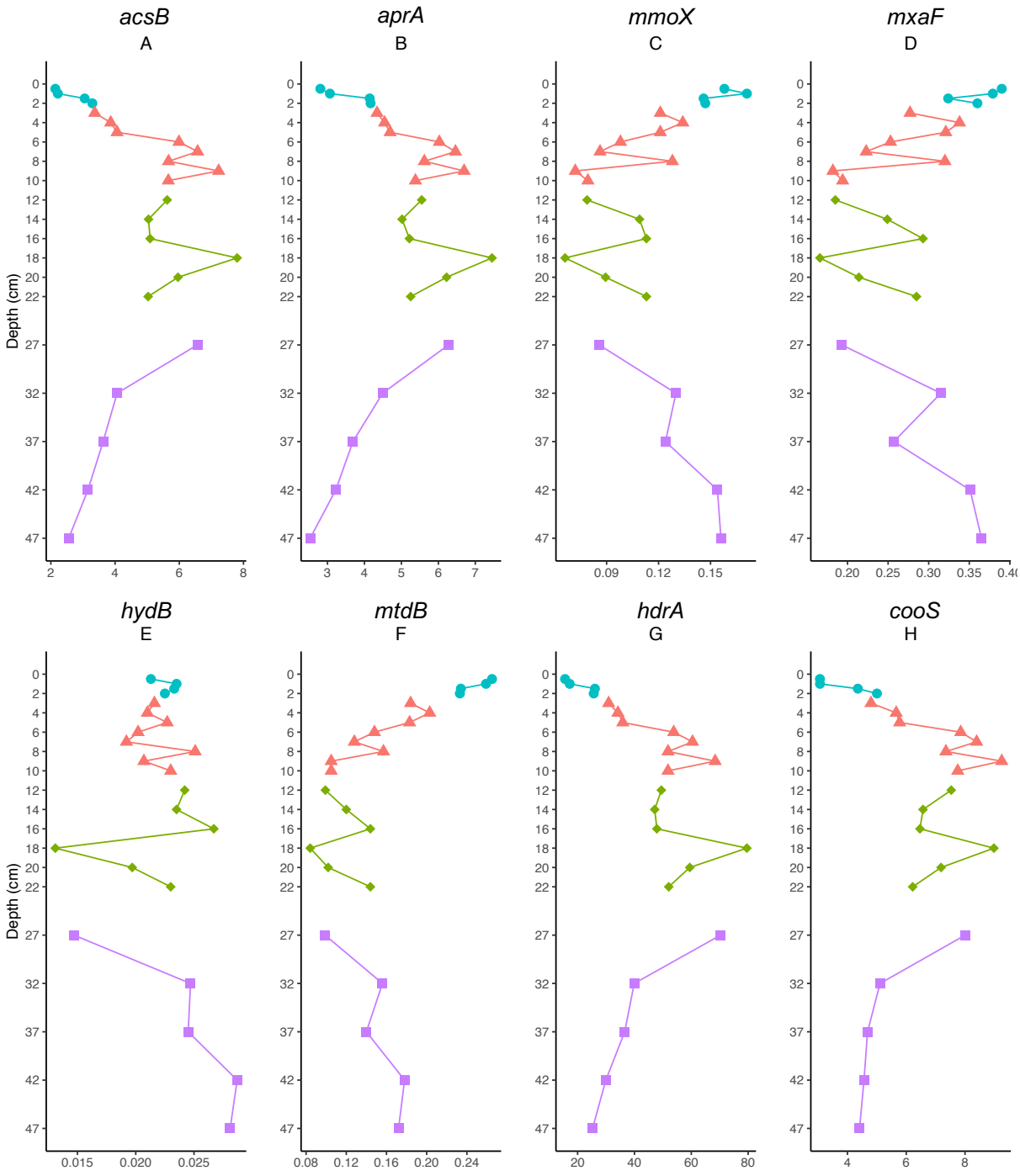


Figure 5.25. Deep Basin Depth Core: Depth profiles of Tax4Fun fractional abundances of a range of target functional genes. (A). *acsB*; (B). *aprA*; (C). *mmoX*; (D). *mxaF*; (E). *hydB*; (F). *mtdB*; (G). *hdrA*; (H). *cooS*.

5.3.7 Geochemical Results

5.3.7.1 *Shallow Basin*

Metal concentrations derived from porewater analysis generally decreased with depth. For instance, manganese (Mn) concentrations decreased from 0.467 mg/L at 0.5 cm to 0.005 mg/L at 35 cm (Figure 5.26A). AOV showed a statistical difference in Mn porewater concentration across the four sampling groups, and Tukey HSD showed a statistical difference in concentration between sample group pairs 1 and all three other groups, and between group 2 – 4 (Appendix E.15A). Porewater iron (Fe) concentrations generally decreased with depth with some sporadic excursions *i.e.* a maximum of 0.373 mg/L at 2 cm and minimums of 0.01 mg/L at 4 and 16 cm, and 0.013 mg/L at 35 cm (Figure 5.26B). AOV and Tukey HSD did not show a statistical difference in porewater Fe concentrations across sample groups (Appendix E.15B). Porewater uranium (U) concentrations were measured in μg/L, and appeared to decrease with depth, with an average concentration of 0.035 ± 0.002 μg/L between 0 – 1.5 cm, 0.013 ± 0.001 μg/L between 2 – 4 cm, peaking at 0.058 μg/L at 5 cm, and then generally decreasing with depth to 0.005 μg/L at 42 cm (Figure 5.26C). AOV showed a statistical difference in U concentration between sample groups, and Tukey HSD showed a statistical difference between sample group pairs 1 and 3, and 1 and 4 (Appendix E.15C).

Anion porewater concentrations had a mixed and complicated depth-related concentration distributions. Porewater nitrate concentrations ranged from 22.27 mg/L at 16 cm and 0 mg/L at 12 cm, with an average concentration of 10.84 ± 0.97 mg/L (Figure 5.27A). AOV and Tukey HSD did not show statistical differences in nitrate concentrations across or between sample groups (Appendix E.15F). Sulphate concentrations did reduce with depth, from 1665.4 ± 77.1 mg/L between 0 – 1.5 cm, to 439.4 ± 19.7 mg/L between 2 – 20 cm, and to 50.1 ± 32 mg/L between 25 – 42 cm (Figure 5.27B). AOV showed a statistical difference in sulphate concentration levels across sample groups, and Tukey HSD showed a statistical difference between sample group 1 and all three other groups, and between pair 2 – 4 (Appendix E.15G).

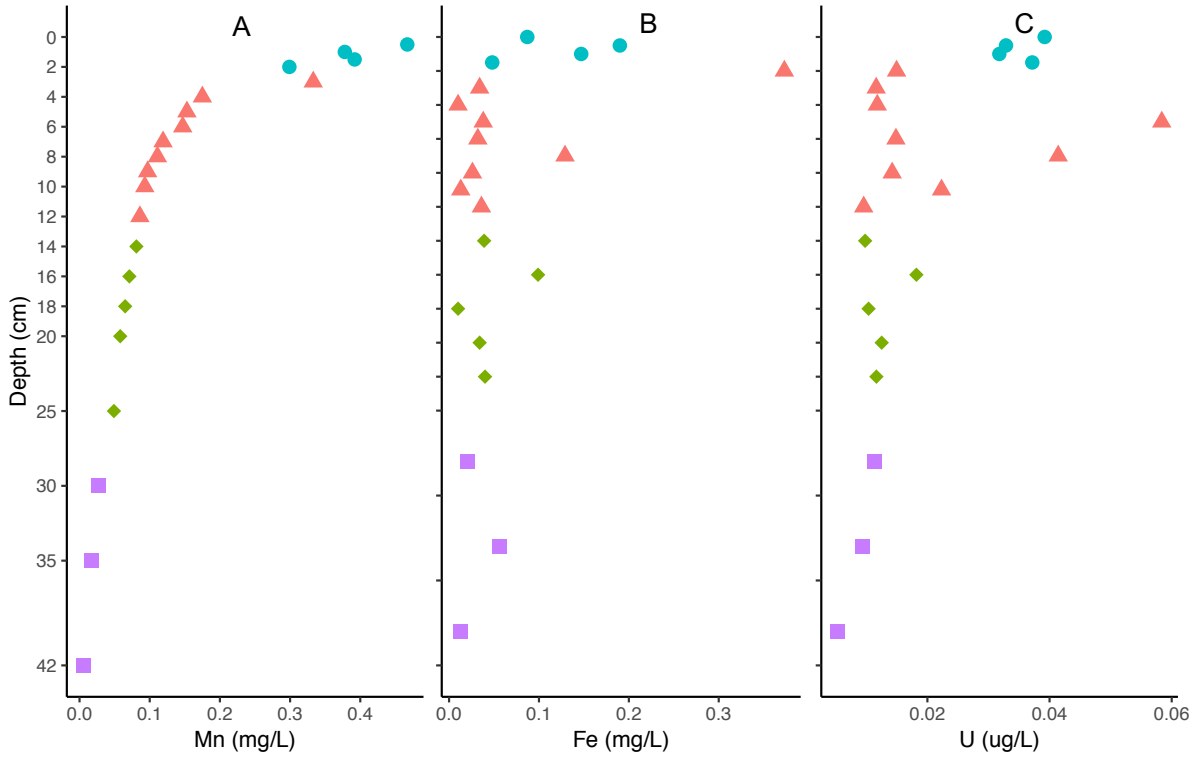


Figure 5.26. Shallow Basin Depth Core: Depth profiles of porewater metal concentrations as a function of depth. (A). Manganese (Mn); (B). Iron (Fe).; (C). Uranium (U).

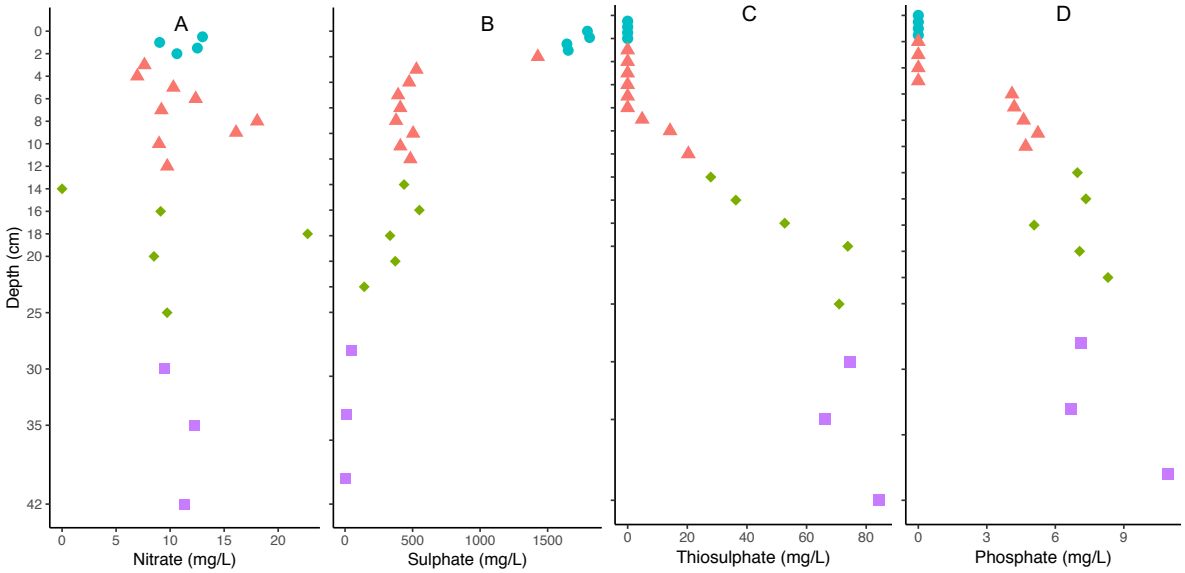


Figure 5.27. Shallow Basin Depth Core: Depth profiles of porewater anion analysis. (A). Nitrate; (B). Sulfate; (C). Thiosulfate; (D). Phosphate.

In contrast to sulfate, thiosulphate concentrations increased with depth, from 0 mg/L between 0 – 8 cm and then increased markedly from 4.86 mg/L at 9 cm to 74.4 mg/L at 30 cm (Figure 5.27C). AOV showed a statistical difference in concentration across all sample groups, and Tukey HSD showed a statistical difference between sample

group pairs 1 – 3, 1 – 4, 2 – 3, and 2 – 4 (Appendix E.15H). Phosphate concentrations similarly increased with depth, from 0 mg/L between 0 – 6 cm, and from 4.1 mg/L at 7 cm to 11 mg/L at 42 cm (Figure 5.27D). AOV showed a statistical difference in concentration across sample groups and Tukey HSD showed a statistical difference between sample group pairs 1 and 3, 1 and 4, 2 and 3, and 2 and 4 (Appendix E.15I).

Solid phase Mn and Fe were normalised to aluminium (Al) to resolve any redox-driven variability with depth (Morford and Emerson, 1999). Mn/Al decreased with depth from a peak of 0.027% at 3 cm to 0.01 at 42 cm, with an average of 0.012 ± 0.0004% between 4 – 42 cm (Figure 5.28A). Fe/Al also decreased with depth albeit more progressively, increasing from 1.04% at 1 cm to 1.11% at 4 cm, and decreased from 1.06% at 5 cm to 0.82% at 42 cm (Figure 5.28B). Due to only one of the samples from sample group 1 (the group sampled at 0.5 cm resolution), AOV was performed on the three groups with sufficient number of samples in (*i.e.*, Groups 2, 3, and 4), which showed a statistical difference across sample resolution groups for both Mn/Al and Fe/Al (Appendix E.15D and 5.16E).

Ferrozine analysis of the 0.5 N HCl extracted solid phase Fe showed that the extractable Fe was present as Fe^{3+} primarily between 0 – 1.5 cm (77.12% at 0.5 cm to 58.5% at 1.5 cm), and at 2 cm, was 65.65% Fe^{2+}, and between 3 – 42 cm, was 98.52 ± 0.38% in the Fe^{2+} species (Figure 5.29). AOV showed a statistical difference in Fe^{2+} fractional abundance across sample resolution groups and Tukey HSD showed a statistical difference between sample group 1 and all other groups (Appendix E.15J).

5.3.7.2 *Deep Basin*

Similar to the Shallow Basin, porewater metal concentrations generally decreased with depth. Mn concentrations decreased relatively quickly, from 1.07 ± 0.1 mg/L between 0 – 5 cm, to 0.692 mg/L at 6 cm, and to 0.531 ± 0.023 mg/L between 7 – 47 cm (Figure 5.30A). AOV showed a statistical difference in porewater Mn concentration across sample groups, and Tukey HSD showed a statistical difference between sample group pairs 1 and all other groups, in addition to group

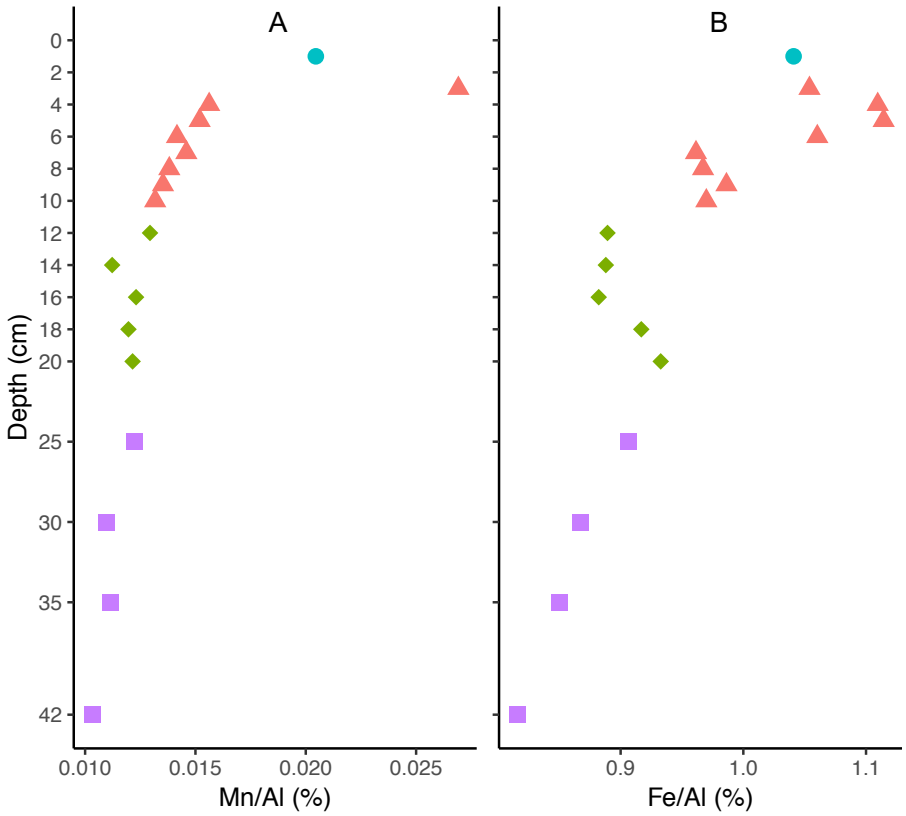


Figure 5.28. Shallow Basin Depth Core: Depth profiles of XRF analysis presented as a percentage of abundance normalised by aluminium (Al) abundance. (A). Mn/Al; (B). Fe/Al.

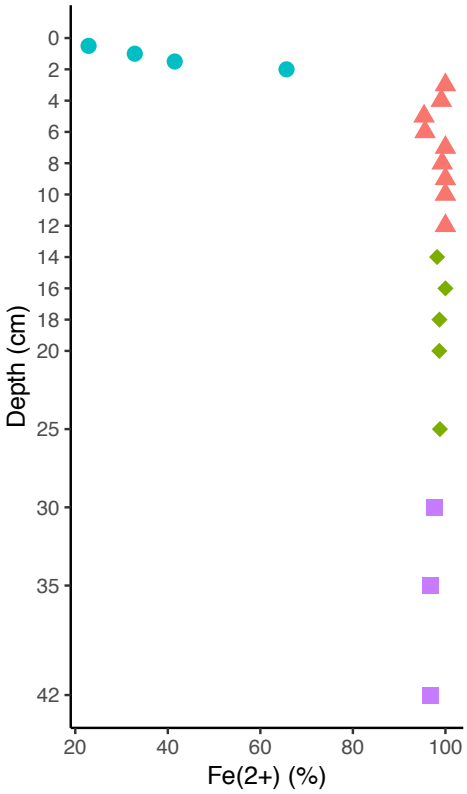


Figure 5.29. Shallow Basin Depth Core: Depth profile of ferrozine analysis showing abundance of Fe^{2+} as a percentage of $Fe^{3+/2+}$.

pair 2 – 4 (Appendix E.16A). Porewater Fe concentrations initially increased with depth and persisted more with depth than Mn. At 0.5 cm, porewater Fe concentration was 0.02 mg/L, 0.005 mg/L at 1 cm, and 0.033 mg/L at 1.5 cm, and this increased to 0.2 mg/L at 4 cm, and between 5 – 22 cm, the average concentration was 0.126 ± 0.01 mg/L, and between 27 – 42 cm, decreased to 0.028 ± 0.013 mg/L (Figure 5.30B). AOV showed a statistical difference in porewater Fe concentrations across all sample groups, and Tukey HSD showed a statistical difference between sample group pairs 1 – 2, 1 – 3, 2 – 4, and 3 – 4 (Appendix E.16B). Porewater U concentration initially decreased with depth, from a peak of 0.055 μg/L at 1 cm to 0.008 μg/L at 6 cm, and average concentration was 0.017 ± 0.004 μg/L between 7 – 47 cm, with a peak of 0.067 μg/L at 37 cm (Figure 5.30C). AOV showed a statistical difference in porewater U concentrations across sample groups, and Tukey HSD showed a statistical difference between sample group pairs 1 – 2, 1 – 3, and 3 – 4 (Appendix E.16C).

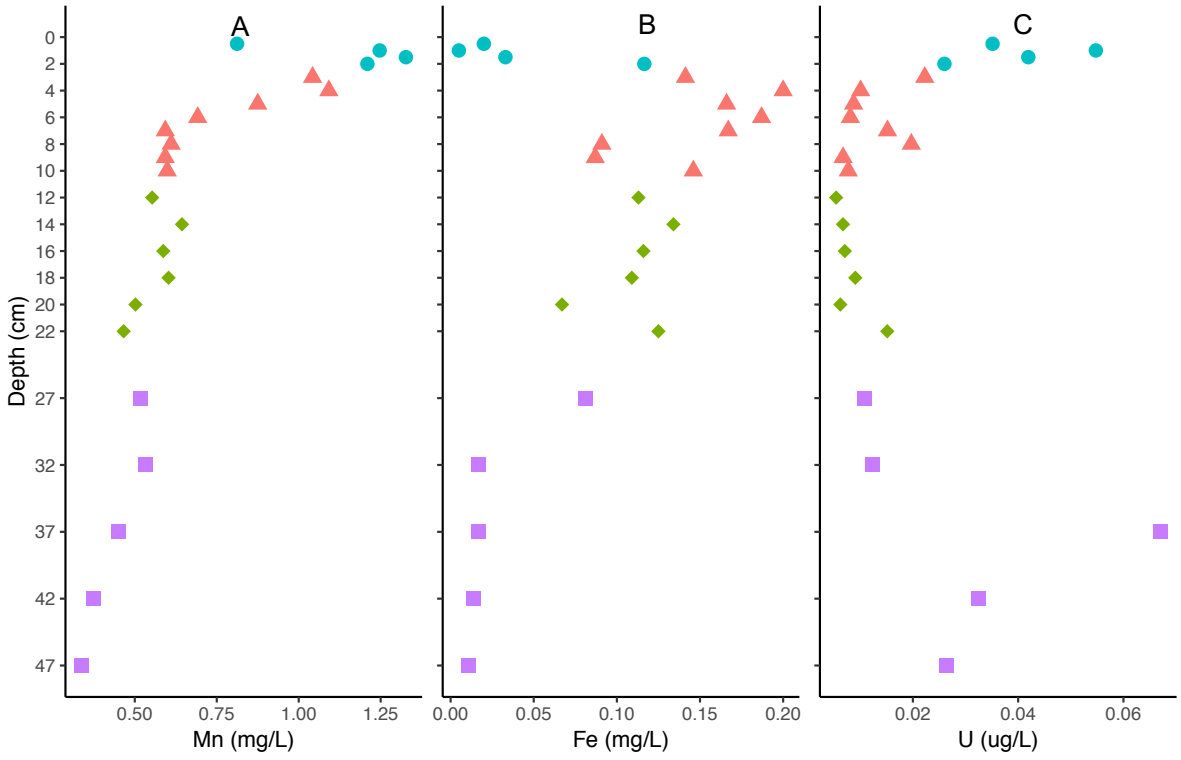


Figure 5.30. Deep Basin Depth Core: Depth profiles of porewater metal concentrations as a function of depth. (A). Mn; (B). Fe.; (C). U.

Anion porewater concentrations appeared to lack depth-related concentration patterns observed in the Shallow Basin depth core. Nitrate porewater concentrations

ranged from 0.15 mg/L at 22 cm to 0 mg/L at 4 cm, averaging 3.09 ± 0.567 mg/L between 0 – 47 cm (Figure 5.31A). AOV and Tukey HSD did not show a statistical difference in nitrate concentrations across or between sample groups (Appendix E.16F). Likewise, sulphate porewater concentrations similarly varied with depth, from 2249.86 mg/L at 22 cm to 342.86 mg/L at 5 cm, averaging 900.1 ± 110.38 mg/L between 0 – 47 cm (Figure 5.31B). AOV and Tukey HSD did not show a statistical difference in porewater sulphate concentrations across or between sample groups (Appendix E.16G). Thiosulphate concentrations were nil for all but one sample, and therefore, that one result was considered unreliable (Figure 5.31C) (Appendix E.16H). Phosphate porewater concentration levels were zero between 0 – 22 cm, and 2.356 ± 0.318 mg/L between 27 – 47 cm (Figure 5.31D), and AOV showed a statistical difference in phosphate concentrations across all sample groups and Tukey HSD showing a statistical difference between sample group pair 4 and all other groups (Appendix E.16I).

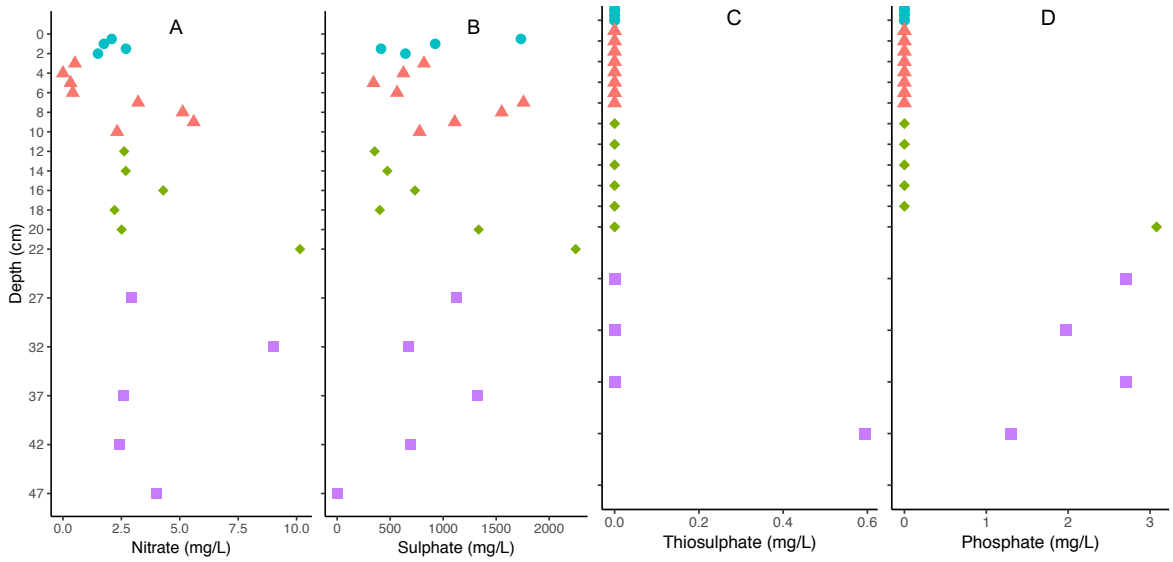


Figure 5.31. Deep Basin Depth Core: Depth profiles of porewater anion analysis. (A). Nitrate; (B). Sulfate; (C). Thiosulfate; (D). Phosphate

Mn/Al decreased with depth from a peak of 0.254% at 1 cm to 0.055 at 22 cm, with an average of 0.095 ± 0.006% between 6 – 42 cm (Figure 5.32A) (no figures were obtained for 47 cm). AOV showed a statistical difference in Mn/Al percentage and Tukey HSD showed a statistical difference between all sample group pairs, with the exception of pair 2 – 3 (Appendix E.16D). Fe/Al decreased with depth, from 1.36% at 1 cm to 0.875% at 42 cm (Figure 5.32B). AOV showed a statistical difference in Fe/Al percentage across all sample groups and Tukey HSD showed a

statistical difference between all sample group pairs except pair 3 – 4 (Appendix E.16E).

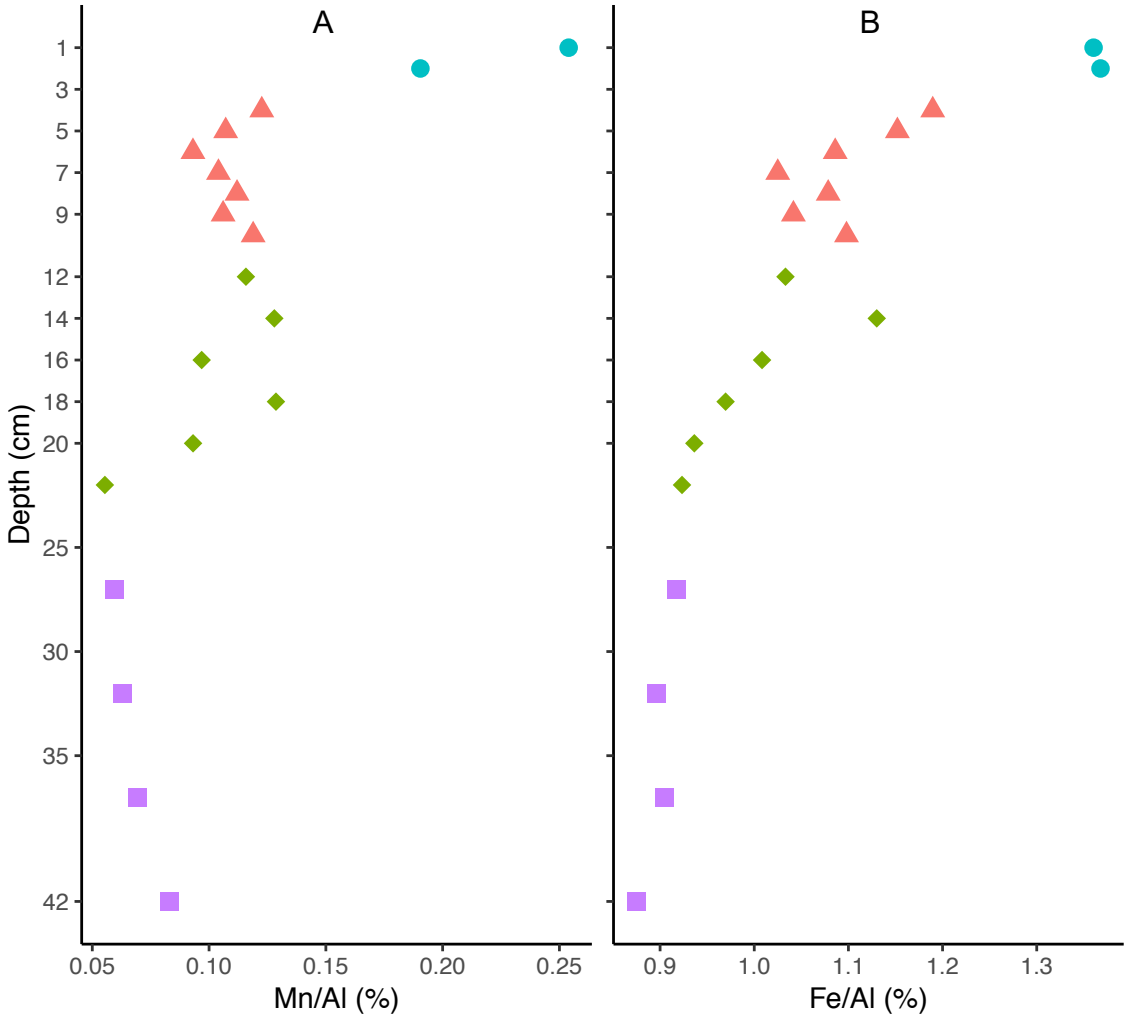


Figure 5.32. Deep Basin Depth Core: Depth profiles of XRF analysis presented as a percentage of abundance normalised by Al abundance. (A). Mn/Al; (B). Fe/Al.

Ferrozine analysis of the 0.5 N HCl extracted solid phase Fe showed that the extractable Fe was present as 80.7% Fe^{3+} at 0.5 cm, 82% Fe^{2+} at 1 cm, 55.7% Fe^{3+} at 1.5 cm, and then an average of 93.27 ± 1.7% Fe^{2+} between 2 – 47 cm (Figure 5.33), with AOV showing a statistical difference in Fe^{2+} percentage across all sample groups and Tukey HSD showing a statistical difference between sample group pair 1 and all other groups (Appendix E.16J).

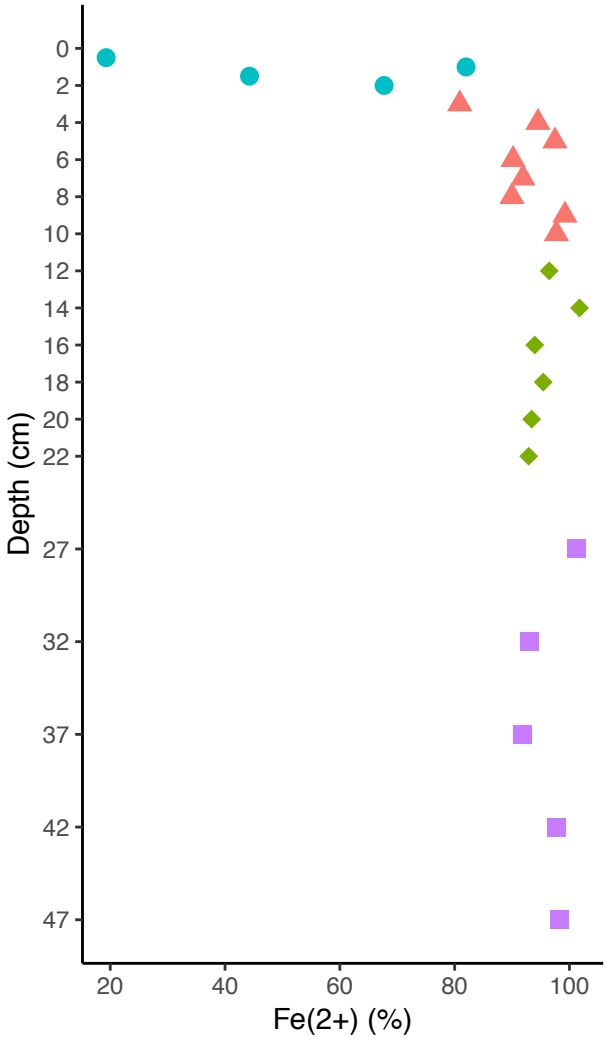


Figure 5.33. Deep Basin Depth Core: Depth profile of ferrozine analysis showing abundance of Fe^{2+} as a percentage of $Fe^{3+/2+}$.

5.3.8 Environmental Fitting

DCA plots of microbial community assemblages between samples for each core were produced as above, with the addition of environmental variable fitting using porewater metal and anion values, and XRF values, to identify any potential correlations with geochemical variables and depth samples. From the Shallow Basin, the porewater Mn and sulfate, and solid phase Mn/Al and Fe/Al concentrations and volumes appeared to correspond with 1 – 1.5 cm samples, while phosphate and thiosulfate levels appeared to correspond with 30 – 42 cm samples (Figure 5.34). At the Deep Basin, porewater Mn and solid phase Mn/Al and Fe/Al values corresponded with the 0.5 – 5 cm samples, porewater nitrate and solid phase values corresponded with the 8 – 42 cm samples, and porewater Fe and U values did not correspond with any specific samples (Figure 5.35).

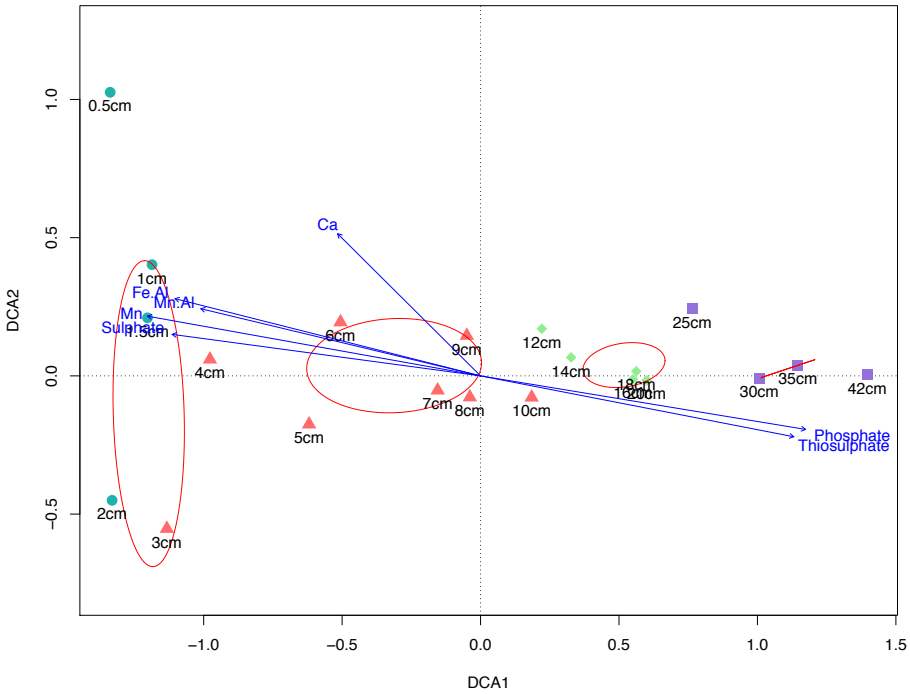


Figure 5.34. Shallow Basin Depth Core: Detrended correspondence analysis (DCA) of Shallow Basin microbial community assemblage with environmental variables vectors overlaid (blue arrows). Length and direction of the arrows relates to correlation between geochemical variable and DCA vectors. Red outline circles denote hierarchical clustering of samples based on centroid dissimilarity.

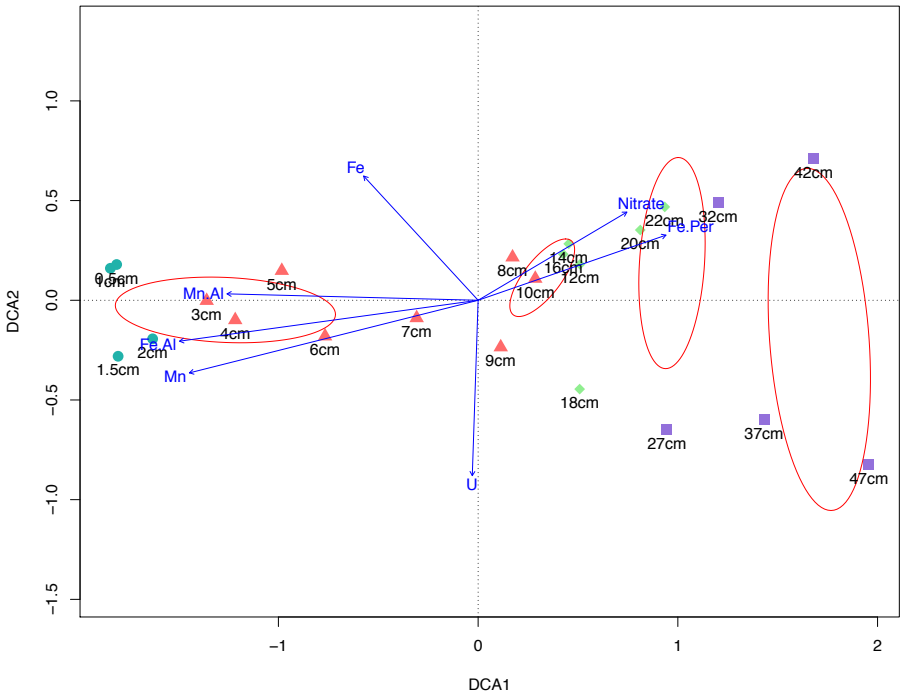


Figure 5.35. Deep Basin Depth Core: Detrended correspondence analysis (DCA) of Deep Basin microbial community assemblage with environmental variables vectors overlaid (blue arrows). Length and direction of the arrows relates to correlation between geochemical variable and DCA vectors. Red outline circles denote hierarchical clustering of samples based on centroid dissimilarity.

5.4 Discussion

The deposition of Sellafield-derived radionuclides (including U, Am, Pu, and Cs) at the Shallow and Deep Basins sites of Loch Etive was the focus of a recent study (Al-Qasmi *et al.*, 2018). The authors of this paper concluded that "the Loch Etive settings develop anoxic sediments which results in the redox cycling of several metals (Mn, Fe, and natural U). However, redox-active artificial radionuclides did not seem to be significantly affected by the redox chemistry in this system." However, in the study of Al-Qasmi *et al.* (2018) only a relatively limited set of geochemical profiles (solid and aqueous phase Mn, Fe and Mn) were used to infer the redox processes occurring. It was anticipated prior to the collection of samples from Loch Etive for this thesis, that depth cores from the loch would exhibit the 'classical' vertical thermodynamically controlled stratification of redox processes (Thauer *et al.*, 1970), with no previous microbiological studies of the loch. As in the previous chapters of this thesis, it was anticipated that depth-resolved characterisation of microbial communities could provide novel insights into the biogeochemical processes occurring at the loch as a function of depth. In the discussion below, it is apparent that the distribution and activity of redox related geochemical processes involving C, N and S cycling are variable and mixed and can be tightly coupled. Furthermore, as with the other sites investigated in this thesis, it is clear that local geographical and hydrological features always impose specific constraints on the disposition of microbial communities and their functions. In many respects, these cores, which were chosen and sampled to provide a classical benchmark for redox succession and its interpretation by community analysis, was the most complex and problematic to interpret, and consequently, it has not been possible to fully reconcile differences between geochemical and taxonomic or inferred functional profiles. This is of interest because the central conclusion of the study of Al-Qasmi *et al.* (2018) was that there was redox cycling of Fe, Mn, and U, but, counterintuitively, not of other redox-active artificial radionuclides in this system.

5.4.1 Comparison of General Characteristics of Microbial Community Structures at the Shallow and Deep Basin

Though the two cores come from the same loch only a few hundred metres apart, the differences in environmental conditions are stark, with the Deep Basin being over 100 metres deeper and anoxic for much longer periods of time. Despite

the differences in environmental conditions, there were still some similarities in the microbial communities between the two sites. For instance, extracted quantities of DNA from both the Shallow and the Deep Basin cores were very similar, ranging from ~ 4 ng/μL to ~ 0.5 ng/μL (Figure 5.1), values which were very similar to those extracted from the Irish Sea Mud Patch (Figure 4.2A). However, while extracted DNA volumes from the Mud Patch did not change as a function of depth, quantities from both Loch Etive cores decreased significantly with depth (Appendix E.1). These data alone indicate considerable changes in microbial activities, probably as a function of carbon inputs and its diagenesis, driving aerobic and anaerobic chemohetero- and chemolithotrophic process. Although depth resolved TOC values were not measured for these cores here or by Al-Qasmi *et al.* (2018), in the Ravenglass and Needle's Eye sites described in previous chapter, DNA quantities had given a significant correlation with TOC values (Figures 3.2 and 4.2), both decreasing with depth. At the Mud Patch, neither DNA nor TOC significantly changed with depth (Figure 4.4). Based on these observations, it is reasonable to infer that TOC probably reduced substantially with depth in both Loch Etive cores, and it is possible to infer that the nature, *i.e.*, bioavailability, of this organic carbon also changed with depth (Parkes *et al.,* 2014). It is worth noting that the age of these sediments has been previously estimated from ^{210}Pb activity measurements (Al-Qasmi *et al.*, 2018) with the 10 cm depths approximately dated to 1986 and 1976 respectively for the shallow and deeper basins. Sedimentation rates appear to be linear in both basins.

From a microbial community structure perspective, both cores exhibited similar patterns of alpha diversity. Species richness was slightly higher in the Deep Basin than the Shallow Basin and did not change as a function of depth in either core (Figures 5.4A and 5.4B), indicating that the cores contained as many unique microbial species at the surface as at the bottom of the cores. Phylogenetic diversity increased in both cores (Figures 5.4C and 5.4D), but only statistically in the Deep Basin (Appendix E.4D), indicating that the microbial species present became more distantly-related at depth, perhaps hinting at increased catabolic functional capacity at depth. Species evenness decreased with depth in both cores, indicating that the communities became more dominated by fewer individual ASVs with depth (Figures 5.4E and 5.4F), statistically more so in the Shallow Basin depth core (Appendix E.4E), despite the fact that the communities were as species rich and more phylogenetically diverse at depth.

Beta diversity of the microbial community assemblages of the two cores resulted in similar depth-related distribution of samples on the ordination plots. The nMDS and DCA plots both showed the two sites had distinct microbial ASV community assemblages, and samples from the two cores were distributed along the axes in the same depth order, generally in the same direction. This indicates that depth was the main variable dictating microbial community assemblage within the cores, and that the microbial assemblages of the two cores were affected in the same way with depth (Figure 5.5). Ordination of the taxonomic classifications of the ASVs again showed the two sites had distinct taxonomic communities, that both sites changed in the same way as a function of depth but with the exception of the Shallow Basin, which after ~ 6 – 9 cm showed the taxonomic community changed with depth differently to the 0.5 – 9 cm of the Shallow Basin and the entire Deep Basin cores (Figure 5.6). This was reflected in the taxonomic barplots, which show a gradual and progressive community assemblage change in the Deep Basin (Figure 5.16) but more of a step change ~ 6 cm in the Shallow Basin (Figure 5.14). PCoAs of individual cores resulted in the 'arch' effect that is common to large, sparse microbial datasets (*i.e.*, datasets that show many ASVs restricted to only a few samples and hence many zeroes in the data table), which is evidence of saturation of the community dissimilarity and is generally indicative of very dissimilar samples at either end of the arch (Morton *et al.*, 2007a). Bray Curtis and weighted UniFrac dissimilarities (which take abundances of ASVs into account) of the Shallow Basin showed general depth-related distribution in an arch (Figure 5.7A) and a 'V' shape (Figure 5.7B) respectively, while unweighted UniFrac (presence/absence of ASVs) did not result in an arch, implying that abundance of the ASVs was a large driver in community dissimilarity (Figure 5.7C). This was further reflected in the pairwise PERMANOVA *p* values, where unweighted UniFrac was the only metric that had a non-significant group pair, groups 3 – 4 (Appendix E.5C). nMDS and DCA plots of the Shallow Basin core showed general distribution of samples in order of depth, but not in the linear fashion observed in the PCoAs (Figure 5.8). The nMDS showed much more heterogeneity in sample depth distribution than either the PCoA and DCA, indicating that, on an intra-core basis, the sample community change was not as linear and progressive as the PCoA implied. This is because the number of dimensions is predetermined before ordination for nMDS, so if two dimensions are chosen, all of the dissimilarity will be displayed, whereas PCoA is a 2D representation of two principal component axes of what could be of many. DCA is known to eliminate the

arch effect of PCoA by detrending (Hill and Gauch, 1980), however, as was demonstrated in the 0.5 – 4 cm samples of the Shallow Basin, as it cannot resolve more than one underlying environmental gradient (Kuczynski *et al.*, 2010). The Deep Basin PCoAs showed an even clearer defined arch effect, including the unweighted UniFrac, implying that the ASV community change as a function of depth was less driven by abundance as in the Shallow Basin, but instead indicates a community that is comprised of more different ASVs at either end of the core (Figure 5.9, Appendix E.6). A cluster of samples between 8 – 18 cm was observed in all three PCoAs. The supposition that the community assemblage is driven by different ASVs at either end of the core, as opposed to abundance-driven at Shallow Basin, corroborates with the increased species richness and phylogenetic diversity observed in the alpha diversity indices for the Deep Basin at depth (Figure 5.4). nMDS of the Deep Basin community assemblage showed depth-related distribution along the nMDS1 axis, with a cluster of samples between 8 – 16 cm, with 20 cm as an outlier (Figure 5.10A), while DCA resulted in a half 'diamond' shape after ~ 10 cm, which is the result of detrending contorting the ordination space, further evidence of greater community dissimilarity between surface and depth samples (Hill and Gauch, 1980).

In addition to beta diversity between the two sites, Gneiss differential abundance (Morton *et al.*, 2017b) also showed distinct microbial ASV communities between the two cores, with some degree of overlap (Figure 5.11). The Loch Etive site was the only site sequenced for this project that resulted in any overlap in ASVs identified between pairs of cores from a given site. Gneiss showed that the whole ASV community balance was more abundant in the Shallow Basin than it was in the Deep Basin (Figure 5.12), with taxonomic proportion plot of the same balance showing that *Clostridiales* ASVs were proportionally more abundant in the Shallow Basin, and *Ignavibacteriales* and *Myxococcales* most abundant in the Deep Basin (Figure 5.13).

Overall, while the Shallow and Deep Basin depth cores appeared to have the same depth-related distribution of community structure and assemblages, there were specific differences between the two that were broadly characterised as the Deep Basin having a more progressive change in microbial community that was driven by the emergence of different ASVs at depth with a cluster of samples between 8 – 20 cm, while the Shallow Basin had more of a step-change around 6 – 9 cm, and that change was driven more by abundance of the ASVs present rather than the

emergence of different ASVs. Taxonomic, functional, and geochemical analysis of each individual core will be discussed further below.

5.4.2 The Shallow Basin Depth Core: Specific Taxonomic, Functional and Geochemical Features

Overall, the taxonomic, functional, and geochemical profiles of the Shallow Basin depth core reveal a core that is putatively sulfate-reducing at the surface, transitioning to fermenting at depth, but with some contradictions, including a lack of observed nitrate reduction despite inferred denitrifying genes, the emergence of sulfur-oxidising taxa and persistence of other aerobic taxa at depth, and a complex sulfur-cycling story pertaining to the reduction of sulfate and introduction of thiosulfate with depth. Porewater U levels decreased from 0.5 – 10 cm as expected from the predicted and observed sulfate-reducing conditions.

'End point' PCR band intensities of the four target genes appeared to show a mixed picture of stratification of aerobic and anaerobic gene abundances. The archaeal *amoA* was very low throughout (Figure 5.2A), indicating a lack of archaeal ammonia oxidisers in this core, which was reflected in the 16S rRNA gene taxonomic analysis. Bacterial *amoA* gene was very abundant (*i.e.*, close to the peak of the band intensity pixel count) from 0.5 – 10 cm, significantly decreased between 12 – 20 cm, and declined to 47 cm (Figure 5.2B, Appendix E.2B), indicating the persistence of bacterial ammonia oxidising taxa until 10 cm, and their re-emergence between 25 – 30 cm. The *nirK* gene appeared to increase between 0.5 – 7 cm, decreased significantly between 12 – 20, and appeared to re-emerge at 42 cm (Figure 5.2C, Appendix E.2C), while the *nirS* decreased progressively with depth (Figure 5.2D), creating a mixed profile of redox conditions at depth.

Several taxa were more abundant between 0.5 – 10 cm that decreased thereafter. *Fusibacter* spp. is a genus capable of chemoorganotrophic fermentation, producing acetate, CO_2, and H_2, and additionally, both chemoheterotrophic anaerobic thiosulfate and elemental sulfur reduction, producing sulfide, but critically, not reducing sulfate or sulfite (Ravot *et al.*, 1999; Ben Hania *et al.*, 2012; Smii *et al.*, 2015; Fadhlaoui *et al.*, 2015; Serrano *et al.*, 2017). With the *Fusibacter* being most abundant in the 0 – 4 cm zone (dropping from 8.6% abundance at 4 cm to 1.3% at 5 cm, and to 0.1% at 7 cm), it is not immediately clear from the 16S rRNA profile alone

whether the *Fusibacter* in the Shallow Basin are fermenting or respiring, though it is worth noting that the thiosulfate concentrations measured in the porewater were zero between 0 – 8 cm (Figure 5.28C), so it is reasonable to deduce that the *Fusibacter* in this core are not respiring using thiosulfate (or perhaps have exhausted all detectable traces of thiosulfate). In contrast to the geochemical and taxonomic mismatch of the *Fusibacter* spp. and thiosulfate concentrations, the *Desulfobulbaceae*, a family generally comprised of incomplete (acetate producing) anaerobic sulfate-reducers (Kuever *et al.*, 2014), were most abundant between 0.5 – 10 cm, which coincided with the reduction in porewater sulfate concentrations observed between 0 – 20 cm, a significant amount occurred between 0 – 3 cm (Figure 5.28B and Appendix E.15G). The *Desulfobulbaceae* and sulfate concentrations had a statistically significant positive correlation (Spearman correlation: $\rho = 0.841$, $p = < 0.001$), implicating the *Desulfobulbaceae* in sulfate reduction at the surface of the core. Other taxa restricted to the upper region of the Shallow Basin depth core included an uncultured γ-*proteobacteria* (Figure 5.14), the representative sequence of which, when subjected to a BLAST analysis of the NCBI nucleotide database, was identical to other uncultured γ-*proteobacteria* sequences, with the first 13 listed of which matched sequences recovered from marine environments, including those contaminated with heavy metals and oil, indicating that this particular sequence is likely one of a relatively ubiquitous marine species. The uncultured δ-*proteobacteria* NB1-j, which has been linked to the *nifD* gene (de Voogd *et al.*, 2015), increased between 0.5 – 9 cm and decreased thereafter (Figure 5.15B), the profile of which matched that of the *nifD* as inferred by Tax4Fun (Figure 5.20F), indicating potential nitrogen-fixation occurring in the upper 9 cm of the core. The *Bathyarchaeia* are widespread in energy-deficient marine subsurface sediments and are acetogenic (He *et al.*, 2016), and contain a homologue of the *mcrA* gene from *Euryarchaeota* and are thought to be have diverse function including a role in methane cycling (Evans *et al.*, 2015). At the Shallow Basin, the *Bathyarchaeia* increased significantly from 0.5 – 5 cm, and persisted throughout the rest of the core (Figure 5.15C, Appendix E.7M), indicating an increased potential for carbon fixation and methane cycling at depth. The *Desulfobacteraceae* Sva0081 group are thought to be sulfate-reducing, H_2-utilising common in coastal marine systems that thrive in mixed redox systems due having to oxidative-stress genes (Mußmann *et al.*, 2015; Probandt *et al.*, 2016; Dyksma *et al.*, 2018), and peaked at 4 cm and then decreased progressively until 30 cm (Figure 5.15D), indicating the presence of hydrogenotrophic sulfate reduction, and a possible

indication chemolithotrophic processes being sustained due to low indigenous carbon levels at the surface of the core. The obligately anaerobic, fermenting, H_2-producing genus *Psychrilyobacter* spp. (Zhao *et al.*, 2009) was most abundant between 6 – 12 cm (Figure 5.15E).

Taxa that increased in abundance with depth after ~ 7 – 10 cm include the *Tepidibacter* spp., an anaerobic and aerotolerant fermenting genus (Urios *et al.*, 2004; Tan *et al.*, 2012), which accounted for as much as 30% of the microbial community assemblage at 30 cm. The *Tepidibacter* spp. were also associated with degradation of complex organic matter in the Irish Sea sediments (O'Reilly *et al.*, 2014). The *Clostridiaceae* genus *Clostridiisalibacter* spp. increased in abundance after 7 cm, and is a halophilic, mesophilic, slightly alkaliphilic, anaerobic, fermenting genus (Liebgott *et al.*, 2008), similar to *Tepidibacter*. Another *Clostridiaceae* genus, *Alkaliphilus* spp., increased in abundance at 7 cm downwards, and is a genus comprised of anaerobic species isolated from a range of relatively extreme environments (including methanogenic and saline environments), and all ferment proteinaceous carbon compounds to acetate, CO_2, H_2, and others, and some species in the genus can use elemental sulfur and thiosulfate as an electron acceptor (Zhilina *et al.*, 2009; Wu *et al.*, 2010; Zakharyuk *et al.*, 2017), whilst others used ferric compounds including Fe^{3+} (Roh *et al.*, 2007; Zhilina *et al.*, 2009; Zakharyuk *et al.*, 2017), and others did not have increased growth with S or Fe compounds, and instead disproportionate or dismutate long chain fatty acids such as crotonate to acetate and butyrate (Coa *et al.*, 2003; Ben Aissa *et al.*, 2015). At 30 cm, the *Tepidibacter* spp. and *Clostridiisalibacter* spp. account for 29% of the total community abundance at that depth. Interestingly, the aerobic, methanesulfate-degrading genus *Filomicrobium* (Henriques and De Marco, 2015) increased with depth (Figure 5.15A). The chemolithotrophic sulfur-oxidising *Thiogranum* spp. (oxidising reduced sulfur compounds) (Mori *et al.*, 2015) increased with depth, peaking at 30 cm (Figure 5.15F), coinciding with the emergence of the thiosulfate concentrations at depth (Figure 5.28C).

Several taxa persisted throughout the core. The strictly anaerobic fermenting *Anaerolineaceae* (Yamada and Sekiguchi, 2018) increased after 2 cm but did not significantly change after 3 cm (Appendix E.7C). The strictly aerobic starch degrading *Sandaracinaceae* (Mohr *et al.*, 2012) persisted until 25 cm before decreasing

(Appendix E.7D). The uncultured *Thermoanaerobaculaceae* subgroup 23 increased after 3 cm and decreased after 20 cm (Appendix E.7E).

Tax4Fun abundance profiles of the 'end point' PCR gene targets did not correlate with the band intensities of the same genes (Figure 5.19), including when data were transformed onto the same scale (*e.g.*, percentage normalised, square-root, or log10 transformed). The *amoA* gene was low and confined mostly to the 1 – 2 cm region (Figure 5.18A). The *nirK* followed the same pattern as the *amoA* and was generally low in abundance as it was in the 'end point' PCR (Figure 5.18B). The *nirS* progressively increased with depth (Figure 5.18C), and the *dsrA* progressively decreased with depth (Figure 5.18D). There were negligible amounts of *mcrA* detected, corroborating with the 'end point' PCR result of the same gene (Figure 5.18E). All other nitrogen cycling genes were inferred to be most abundant between 0.5 – 10 cm, peaking at ~ 4 cm (Figure 5.20). Methane-oxidising genes were low and appeared to be most abundant at ~ 25 – 42 cm (Figures 5.21C, 5.21D, 5.21F). Anaerobic carbon fixating genes *acsB* and *cooS* were most abundant between 0.5 – 12 cm, peaking at 4 cm (Figures 5.21A and 5.21H), and the dissimilatory sulfate reduction pathway *aprA* and methane-cycling *hdrA* also peaked at 4 cm (Figures 5.21B and 5.21G).

Taxonomic and functional gene profiles of the Shallow Basin depict a core that is characterised not by the normally-expected 'classical' thermodynamically controlled stratification of redox processes, but by a transition from sulfate-reducing conditions at surface through to fermenting conditions at depth, and somewhat contradicting the persistence of several strictly aerobic taxa throughout with the emergence of sulfur-oxidising taxa at depth. This heterogenous distribution of aerobic and anaerobic conditions implied by taxonomy and functional gene abundance is intriguing as this pattern ostensibly seems to contradict the measured geochemical profiles of solid phase and dissolved Mn and Fe. From similar 'metals' profiles it was previously concluded that the Loch Etive sediments are stably anaerobic immediately below the surface (Al Qasmi *et al.*, 2018). In the classical model of redox succession, solid phase oxidised iron and Mn are deposited into the sediment and reduced to release dissolved Mn^{2+} and Fe^{2+}. The diffusion of these reduced soluble metals upwards with subsequent re-oxidation and precipitation leads to the surface being a sink for both Mn and Fe. This is why solid phase Mn and to some extent Fe is depleted with depth as observed in Figure 5.29. However, iron is slightly different.

Unlike Mn^{2+}, Fe^{2+} more readily reacts with sulfide produced from sulfate reduction (Suess, 1978) to form mono and disulfide mineral precipitates. This reactivity, in addition to iron's different redox chemistry, could partly explain persistence of solid phase iron at depth. However, in the Shallow Basin depth core there is an intriguing secondary sub-peak of solid phase iron at around 20 cm which, based on the ferrozine analysis of HCl extracts (Figure 5.29B), likely reflects the presence of acid volatile Fe^{2+} monosulfides. Furthermore, this sub-peak of iron sulfide coincides with a similar sub-peak in sulfate at the same depth which may imply an active iron and sulphur cycle at these depths (Figure 5.28B), *i.e.*, involving oxidative as well as reductive reactions. Certainly, the relationship between sulfate (SO_4^{2-}) and thiosulfate ($S_2O_3^{2-}$) porewater concentration profiles observed at the Shallow Basin is supportive of a subsurface sulfur cycling. Sulfate is clearly removed by reduction to H_2S throughout the top 25 cm of the Shallow Basin depth core and the taxonomic depth profiles of the *Desulfobulbaceae* (Figure 5.14) and *Desulfobacteraceae* (Figure 5.15D) are consistent with this pattern. Culture studies of the *Desulfobulbaceae* indicate that the genus reduces sulfate only as far as sulfide (Lien *et al.*, 1998). However, coinciding with the lower section of the sulfate reduction zone and the sub-peaks in iron and sulfate is a corresponding increase of thiosulfate with depth. Thiosulfate is known to be generated in anoxic environments by what can be considered as an 'inorganic fermentation' process (see equations) involving the disproportionation of elemental sulfur (Finster *et al.*, 2008).

$$3S^0 + 3H_2O \leftrightarrow 2S^{2-} + SO_3^{2-} + 6H^+$$

$$S_2O_3^{2-} \leftrightarrow S^0 + SO_3^{2-} \ (\Delta G = -26.7\ kJ\ mol^{-1})$$

The first equation shows the biotic disproportionation of elemental sulfur to sulfite (SO_3^{2-}) performed by microorganisms, while the second shows the reaction of sulfite and elemental sulfur to produce thiosulfate (Thamdrup *et al.*, 1994). Critically, for this process to be operating in the Shallow Basin core sediment there must be a mechanism for oxidative processes facilitating cycling of sulfate *via* generation of sulfur (not measured) which may be the key characteristic of the Shallow Basin from a functional perspective. The emergence of the *Thiogranum* spp. which oxidises a range of reduced sulfur compounds including thiosulfate, sulfite, elemental sulfur, sulfide, and tetrathionate (Mori *et al.*, 2015) (Figure 5.15F) statistically correlated with the observed increase in porewater thiosulfate concentrations (Figure 5.28C) (Spearman correlation ρ value = 0.9, *p* value = < 0.001). The role of *Thiogranum* spp.

in this core in relation to the generation of thiosulfate is unclear, but it is worth noting that *Thiogranum* spp. can only oxidise thiosulfate in the presence of oxygen (Mori *et al.*, 2015).

An alternative method for thiosulfate generation at depth in the Shallow Basin is anoxic sulfide oxidation. Though H_2S will readily oxidise in the presence of O_2, oxidation of sulfide predominately to thiosulfate and also sulfate has been documented in anoxic waters (Tuttle and Jannasch, 1973; Jørgensen *et al*., 1991). Bacteria linking denitrification, specifically nitrate reduction, to sulfur oxidation were postulated to be the cause for these observations. Subsequent work confirmed the role of nitrate reduction in sulfide oxidation in anaerobic waters (Wang *et al*., 2005 Beristain Cardoso *et al*., 2006), and also nitrite reduction (Mahmood *et al*., 2007). Porewater nitrate concentrations did not have any pattern with depth in this core (Figure 5.27A), potentially precluding this mechanism of thiosulfate generation. However, while 'end point' PCR denitrification gene profiles were inverse to thiosulfate concentrations (Figures 5.2), and Tax4Fun gene profiles of most denitrification and nitrogen-cycling genes also generally decreased with depth (Figures 5.18 and 5.20), Tax4Fun inferred an increase in the *nirS* gene with depth (Figure 5.18C). This also coincided with an increase in inferred methane oxidation genes, *mmoX* and *mxaF* (Figures 5.21C and 5.21D). Anaerobic methane oxidation coupled to nitrite reduction has been shown to produce excess free O_2 (Ettwig *et al*., 2010), and this excess O_2 may be oxidising any sulfite or sulfide to thiosulfate.

Thiosulfate can also be generated in anoxic waters when coupled to the reduction of Mn^{4+} oxides and the disproportionation of S° (Yao and Millero, 1995; Yao and Millero, 1996; Böttcher and Thamdrup, 2001). An isolate from the genus *Sulfurimonas* was shown to be able to link Mn^{4+} reduction to the oxidation of H_2S to generate thiosulfate in sulfidic environments (Henkel *et al*., 2019). This genus was identified in the Shallow Basin taxonomic analysis, but only in single digit copies in two samples. In the Shallow Basin depth core, solid phase Mn generally decreased with depth (Figure 5.28A) as would be expected in an anoxic environment. The lack of porewater Mn (soluble Mn^{2+}), the expected product of Mn^{4+}-linked thiosulfate generation, at depth would appear to preclude this, but observed porewater Mn accumulation at surface discussed above means it remains a possible mechanism.

Abiotic reduction of U^{6+} to U^{4+} by HS^- results in the production of elemental S (Hua *et al*., 2006). Porewater U concentrations were inverse with thiosulfate

concentrations in the Shallow Basin core (Figures 5.26C and 5.27C). Complete reduction of the sulfate at surface (Figure 5.27B) to sulfide may be abiotically reducing Mn and U to produce thiosulfate and elemental sulfur respectively, and the elemental sulfur may be undergoing disproportionation with any sulfite to also produce thiosulfate.

Further work at elucidating the exact mechanism by which thiosulfate accumulates at depth and how this relates to sulfate reduction observed from the surface downwards is required. A critical question is how might elemental sulfur required for thiosulfate generation is generated. One suggestion might be the observed 16-month overflowing of the sill at Bonawe that introduces oxygenated water at depth in the loch (Edwards and Edelsten, 1977; Ridgway and Price, 1987) leading to the occasional turnover of sediments and introduction of more oxidising chemical species at depth.

5.4.3 The Deep Basin Depth Core: Specific Taxonomic, Functional and Geochemical Features

Where the Shallow Basin had a distinct separation between the overall inferred biogeochemical function between 0.5 ~ 7 cm and ~ 8 – 42 cm, the Deep Basin exhibited a more progressive and less clearly-defined separation of taxonomy and function with depth and contained a similar mix of aerobic and anaerobic taxa and genes throughout the core. As with the Shallow Basin, there was no evidence of microbial taxonomic of functional stratification with depth that followed the usually-expected vertical profile of thermodynamically controlled redox succession (Thauer *et al.*, 1977) which was so obvious in Needles Eye. In some respects, as discussed below, the geochemical and microbial profiles are even more difficult to reconcile with each other than for the Shallow Basin depth core, although it is contradictions within the geochemical profiles which are most perplexing.

'End point' PCR initially indicated that the core was aerobic and denitrifying from 0.5 to ~ 10 cm, based on the band intensities of bacterial *amoA*, *nirK*, and *nirS* genes being statistically highest in this region (Figure 5.3, Appendix E.3).

Many of the most abundant taxa in the core had a gradual and progressive change in abundance as a function of depth (Figure 5.16). Taxa that decreased gradually with depth included the aerobic starch degrading *Sandaracinaceae* (Mohr

et al., 2012) (Appendix E.8B) and the *Bacteroidetes* BD2-2, which was associated with heterotrophs, sulfate-reducers, and archaeal methanotrophs (Trembath-Reichert *et al.*, 2016) (Appendix E.8C). The sulfate-reducing *Desulfobulbaceae* (Kuever, 2014) declined rapidly from 4 cm (Figure 5.17, Appendix E.8K). The sulfate-reducing *Desulfobacteraceae* Sva0081 described above (Mußmann *et al.*, 2015; Probandt *et al.*, 2016; Dyksma *et al.*, 2018) decreased gradually from 2 – 37 cm (Figure 5.17E, Appendix E.8O).

Taxa that appeared to increase until the middle of core and decrease thereafter appeared to be mostly comprised of δ-*proteobacteria*, including the syntrophic and strictly anaerobic *Syntrophobacteraceae* (Kuever, 2014b) although this did not have a significant difference in abundance (Appendix E.8G), the δ-*proteobacteria* NB1-j (Figure 5.17B, Appendix E.8L), linked to the *nifD* gene (de Voogd *et al.*, 2015), and the *Desulfobacteraceae* SEEP-SRB1 group, commonly associated with sulfate reduction and anaerobic methane oxidation (Timmers *et al.*, 2015) increased between 0.5 – 8 cm, decreased gradually to 32 cm, and remained low thereafter (Figure 3.17F, Appendix E.8P). One exception from the δ-*proteobacteria* was the an anaerobic and aerotolerant fermenting genus *Tepidibacter* spp., (Urios *et al.*, 2004; Tan *et al.*, 2012), which was most abundant between ~ 7 – 20 cm (Figure 5.16, Appendix E.8J).

Taxa that increased at depth in the core included the strictly anaerobic fermenting *Anaerolineaceae* (Yamada *et al.*, 2006), although this increased after 2 cm and did not statistically increase after that (Appendix E.8A). An uncultured *Clostridiaceae*, that when the representative sequence of the ASV was entered into BLAST, returned as 98% match to the anaerobic *Clostridium* spp. (Collins *et al.*, 1994) sequence from a marine sponge. The *Bathyarchaeia*, described above as a methane cycling class, increased with depth and was most abundant after 18 cm (Appendix E.8F). Both the uncultured *Ignavibacteriales* PHOS-HE35 and the *Ignavibacterium* spp. generally increased with depth, although in the case in the *Ignavibacterium* spp., this was in the form of two separate peaks at 16 cm and 42 cm (Figure 5.17). The *Ignavibacterium* is a strictly anaerobic chemoheterotrophic genus (Iino *et al.*, 2010), and its closest relative is capable of anaerobic respiration using nitrate, iron, and arsenic as TEAs (Podosokorskaya *et al.*, 2012). The strictly anaerobic sulfate-reducing *Desulfatiglans* spp., which is a complete oxidiser that uses sulfate exclusively (Suzuki *et al.*, 2014), increased with depth (Figure 5.17C).

Tax4Fun resulted of the 'end point' PCR target genes showed *amoA* and *nirK* relatively low in abundance between 0 – 22 cm, and then being at their most abundant at 47 cm (Figures 5.23A and 5.23B), indicating a lack of these functions until 47 cm. The *nirS* gene increased progressively with depth (Figure 5.23C). The *dsrA* gene increased between 0 – 18 cm and decreasing thereafter (Figure 5.23D) and no *mcrA* was inferred, as expect based on the Shallow Basin, 'end point' PCRs, and the Ravenglass and Mud Patch cores. The denitrifying *narG*, *napA*, and *norB* all increased between 0.5 – 18 cm, and decreased after (Figures 5.25A, 5.25B, and 5.25C), while the final denitrification *nosZ* gene persisted from 0.5 cm to 27 cm (Figure 5.25D). The DNRA nitrite reductase *nrfA* increased from 0.5 – 5 cm, decreased at 6 cm, and increased gradually throughout the rest of the core (Figure 5.25E). The nitrogen-fixing *nifD* increased from 0.5 – 18 cm and decreased thereafter (Figure 5.25F). The anaerobic carbon-fixing *acsB* and *cooS*, the dissimilatory sulfate reduction pathway *aprA*, and the methanogenic/methanotrophic *hdrA* all increased from 0.5 to 18 cm then decreased to 47 cm (Figures 5.26A, 5.26H, 5.26B, and 5.26G respectively). Methane-oxidising genes including *mmoX*, *mxaF*, and *mtdB* all showed an inverse pattern to the anaerobic genes, decreasing from 0.5 to 18 cm, then increasing thereafter (Figures 5.26C, 5.26D, and 5.26F respectively).

Based on the microbial taxonomic and functional gene profiles described above, it might be expected that geochemical profiles for the Deep Basin depth core would follow a progressive transition with depth with evidence of denitrification, iron reduction, and sulfate reduction from 0.5 – 27 cm. With increasing depth, highly reducing chemoorganotrophic taxa and function became more dominant. However, as with the Shallow Basin depth core with depth, the increase in several aerobic genes and taxa (*amoA*, *mmoX*), including for the Deep Basin core the chemolithoautotrophic *Bathyarchaeia* (Evans *et al.*, 2015), implies the presence of more oxidizing conditions at depth. In fact, while the solid phase Mn and Fe and porewater Mn, Fe, and U levels reduced with depth (Figures 5.33 and 5.31 respectively), consistent with successive reduction of these metals as described for the Shallow Basin depth core and implying the reduction of these elements, porewater nitrate and even in this case sulfate levels did not decrease with depth (although the values shown in Figure 5.32 are very variable from one depth to another). Interestingly, with respect to this lack of sulfate reduction the thiosulfate accumulation observed at the Shallow Basin was not observed at the Deep Basin

insofar that the thiosulfate was apparently absent in the Deep Basin porewater, and the sulfate did not appear to be reduced in the Deep Basin as it was in the Shallow Basin. The potential complex sulfur cycling that characterises the Shallow Basin core appears to be absent in the Deep Basin core. The lack of sulfate reduction in the porewater seemingly contradicts the presence of known-sulfate-reducing δ-*proteobacteria* observed in the 16S rRNA gene taxonomy profile at the surface of the core. The lack of thiosulfate accumulation may therefore be due to a lack of sulfide (a product of sulfate reduction) available for disproportionation (Jørgensen and Bak, 1991; Thamdrup *et al.*, 1994).

5.4.4 The Fate of Radionuclides in the Loch Etive Cores

Based on the microbial taxonomic and functional profiles generated here, it is possible to hypothesise about the transport and redox fate of actinide radionuclides in the two Loch Etive cores, and to compare these inferences with those postulated in Al-Qasmi *et al*. (2018), but without any solid phase XRF data or redox species HERFD-XANES data, it is not possible to confirm these hypotheses directly. Furthermore, neither of these cores appears to follow the normally expected thermodynamically controlled vertical stratification of redox couples (Thauer *et al*., 1977), instead demonstrating evidence of complex sulfur cycling.

The Shallow Basin depth core appears to be sulfate reducing from surface to ~ 10 cm, and in this region, any actinides present would be expected to be reduced to at least their tetravalent, and insoluble, species. The presence of the sulfate-reducing bacteria and inferred abundance of the *dsrA* gene was highest in the upper 10 cm (Figures 5.15 and 5.18). This corroborated with soluble porewater U measured in this core in this project which was generally confined to the upper 10 cm of the core, reducing with depth (Figure 5.26C), suggesting reduction of U^{6+} in this potentially sulfidic zone. This was further evidenced by the reduction in porewater Mn, Fe, and SO_4^{2-} and ferrozine analysis indicating that Fe in the core was predominantly Fe^{2+} from ~ 2 cm downwards. Al-Qasmi *et al*. (2018) also suggested that U was reduced after 10 cm. Either microbially-mediated bioreduction of U may be occurring here, or abiotic reduction *via* the by-products of other microbially-mediated reduction processes expected to be co-occurring in this zone, *e.g.*, Mn/Fe^{2+}, sulfite, sulfide.

In the lower 12 – 42 cm region of the core, the microbial taxonomic and functional profiles become less definitive than the upper 10 cm region. A heterogenous mix of anaerobic fermenting and aerobic oxidising taxa emerge, and this zone is defined by the generation of thiosulfate at depth – a process normally associated with oxidation of sulfide and thus the presence of O_2. Based on the decreased abundance of known sulfate-reducing taxa and genes in this lower region, it can be inferred that sulfate has been reduced to either sulfite or sulfide. Sulfide has been shown to reduce U^{6+} to generate U^{4+} and elemental sulfur (Hua *et al*., 2006). The emergence of fermenting bacteria in this lower zone potentially indicates the lack of available respiratory terminal electron acceptors (TEAs), which was partly corroborated by geochemical analysis indicating relatively low levels of solid phase Mn and Fe, porewater sulfate, and ferrozine analysis which showed nearly all the Fe in the core to be reduced Fe^{2+}. Therefore, it is very unlikely that any actinides in this part of the core will be in the oxidised, soluble, species, and will either be directly reduced by bacteria, potentially by the *Clostridiaceae* species abundant in this region, which are relatively closely related to the U-reducing *Clostridia* spp. (Gao and Francis, 2008), or abiotically by other microbially-reduced species.

The interpretation of U mobility is broadly similar between this analysis and that of Al-Qasmi *et al*., in that it is anticipated actinides present will be reduced after 10 cm. A lack of Pu/Am accumulation at depth based on historical Sellafield-derived releases was not observed as expected by Al-Qasmi *et al*. and thus inferred that Pu is being further reduced to the mobile and soluble Pu^{3+} species linked to sulfate reduction at depths > 10 cm. Microbial profiles do not corroborate with this hypothesis, which showed the sulfate-reducing taxa and genes confined almost exclusively to the 0 – 10 cm region. Microbial profiles indicate that this region of the core is potentially very reducing and characterised by a lack of available TEAs for respiration, hence the emergence of fermenting bacteria. Pu^{4+} may being reduced directly during respiration to Pu^{3+}, though Pu^{4+} has been shown to be refractory and not readily reduced to Pu^{3+} in anaerobic conditions (Kimber *et al*., 2012), and Pu^{3+} generally oxidises back to Pu^{4+} spontaneously (Rusin *et al*., 1994). Mixing of Sellafield-derived Pu in the Irish Sea Mud Patch as described in Chapter 4 and in Al-Qasmi *et al*. is likely acting as a constant source of Pu, Am, Cs, and U from Sellafield. The homogenous Pu/Am profile in the Shallow Basin observed by Al-Qasmi *et al*. therefore potentially merely reflects a constant input of refractory Pu^{4+} into the Loch.

Making direct inferences about the transport and fate of redox-active radionuclides and actinides in the Deep Basin depth core based on microbial profiles generated was more complicated than for Shallow Basin. The Deep Basin is generally much more anoxic and for greater periods of time than the Shallow Basin (Edwards and Edelsten, 1977), so superficially one would expect to find evidence for reducing and fermentative conditions. This was generally the case, with numerous fermenting taxa found at depth in the Deep Basin core. Numerous putatively sulfate-reducing δ-*proteobacteria* were found throughout the core, with one *Desulfobulbaceae* confined to the upper ~ 6 cm, one *Desulfatiglans* spp. persistent from ~ 12 cm downwards, and the *Desulfobacteraceae* Sva0081 and SEEP-SRB1 persisting until ~ 30 and 37 cm respectively (Figure 5.17). Therefore, based on these microbial profiles, it could reasonably be inferred that with the presence of sulfate-reducing and fermenting bacteria, actinides present would probably be in their tetravalent species (either by direct enzymatic bioreduction of abiotic reduction). However, these sulfate-reducing bacteria profiles did not corroborate with porewater sulfate concentrations, which did not change with depth, and furthermore, Al-Qasmi *et al*. did not find evidence of sulfate nor U reduction in this core.

5.5 References

Al-Qasmi H, Law GTW, Fifield LK, Howe JA, Brand T, Cowie GL, Law KA, Livens FR (2018) Deposition of artificial radionuclides in sediments of Loch Etive, Scotland. Journal of Environmental Radioactivity, 187: 45–52

Ansell AD (1974) Sedimentation of Organic Detritus in Lochs Etive and Creran, Argyll, Scotland. Marine Biology, 27: 263–273

Asshauer KP, Wemheuer B, Daniel R, Meinicke P (2015) Tax4Fun: predicting functional profiles from metagenomics 16S rRNA data. Bioinformatics, 31(7): 2882–2884

Ben Aissa F, Postec A, Erauso G, Payri C, Pelletier B, Hamdi M, Fardeau M-L, Olliver B (2015) Characterization of *Alkaliphilus hydrothermalis* sp. nov., a novel alkaliphilic anaerobic bacterium, isolated from a carbonaceaous chimney of the Prony hydrothermal field, New Caledonia. Extremophiles, 19: 183–188

Ben Hania W, Fraj B, Postec A, Fadhlaoui K, Hamdi M, Olliver B, Fardeau M-L (2012) *Fusibacter tunisiensis* sp. nov., isolated from an anaerobic reactor used to treat olive-mill wastewater. International Journal of Systematic and Evolutionary Microbiology, 62: 1365–1368

Beristain Cardoso R, Sierra-Alvarez R, Rowlette P, Razo Flores E, Gómez J, Field JA (2006) Sulfide Oxidation Under Chemolithoautotrophic Denitrifying Conditions. Biotechnology and Bioengineering, 95(6): https://doi.org/10.1002/bit.21084

Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso JG (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome, 6(90): https://doi.org/10.1186/s40168-018-0470-z

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 13: 581-583

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. The ISME Journal, 11: 2639–2643

Cao X, Liu X, Dong X (2003) *Alkaliphilus crotonatoxidans* sp. nov., a strictly anaerobic, crotonate-dismutating bacterium isolated from a methanogenic environment. International Journal of Systematic and Evolutionary Microbiology, 53: 971–975

Capone DG, Kiene RP (1988) Comparison of microbial dynamics in marine and freshwater sediments: Contrasts in anaerobic carbon metabolism. Limnology and Oceanography, 33(4): 725–749

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Gonzalez Pena A, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7(5): 335-336

Coûteaux M-M, Mousseau M, Célérier M-L, Bottner P (1991) Increased atmosphereic CO_2 and litter quality: decomposition of sweet chesnut leaf litter with animal food webs of different complexitities. OIKOS, 61: 54–64

de Hass H, van Weering TCE, de Stiger H (2002) Organic carbon in shelf seas: sinks or sources, processes and products. Continental Shelf Research, 22: 691–717

Edwards A, Edelsten DJ (1977) Deep Water Renewal of Loch Etive: A Three Basin Scottish Fjord. Estuarine and Coastal Marine Science, 5: 575–595

Ettwig KF, Butler MK, Le Paslier D, Pelletier E, Mangenot S, Kuypers MMM, Schreiber F, Dutilh BE, Zedelius J, de Beer D, Gloerich J, Wessels HJCT, van Alen T, Leusken F, Wu ML, van de Pas-Schoonen KT, Op den Camp HJM, Janssen-Megens EM, Francoijs K-J, Stunnenberg H, Weissenbach J, Jetten MSM, Strous M (2010) Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature, 464: 543–548

Evans PN, Parks DH, Chadwick GL, Robbins SJ, Orphan VJ, Golding SD, Tyson GW (2015) Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. Science, 350: 434–438

Fadhlaoui K, Ben Hania W, Postec A, Fauque G, Hamdi M, Olliver B, Fardeau M-L (2015) *Fusibacter fontis* sp. nov., a sulfur-reducing, anaerobic bacterium isolated from a mesothermic Tunisian spring. International Journal of Systematic and Evolutionary Microbiology, 65: 3501–3506

Finster K (2008) Microbiological disproportionation of inorganic sulfur compounds. Journal of Sulfur Chemistry, 29(3-4): 281-292, DOI: 10.1080/17415990802105770

Henkel JV, Dellwig O, Pollehne F, Herlemann DPR, Leipe T, Schulz-Vogt HN (2019) A bacterial isolate from the Black Sea oxidizes sulfide with manganese(IV) oxide. PNAS: DOI:10.1073/pnas.1906000116

Henriques AC, De Marco P (2015) Complete Genome Sequences of Two Strains of "*Candidatus* Filomicrobium marinum," a Methanesulfonate-Degrading Species. Genome Announc 3(3):e00160-15. doi:10.1128/genomeA.00160-15

Hua B, Xu H, Terry J, Deng B (2006) Kinetics of Uranium(VI) Reduction by Hydrogen Sulfide in Anoxic Aqueous Systems. Environmental Science and Technology, 40: 4666–4671

Iino T, Mori K, Uchino Y, Nakagawa T, Harayama S, Suzuki K-I (2010) *Ignavibacterium album* gen. nov., sp. nov., a moderately thermophilic anaerobic bacterium isolated from microbial mats at a terrestrial hot spring and proposal of *Ignavibacteria* classis nov., for a novel lineage at the periphery of green sulfur bacteria. International Journal of Systematic and Evolutionary Microbiology, 60: 1376–1382

Jørgensen BB, Bak F (1991) Pathways and Microbiology of Thiosulfate Transformations and Sulfate Reduction in a Marine Sediment (Kattegat, Denmark). Applied and Environmental Microbiology, 57(3): 847–856

Jørgensen BB, Fossing H, Wirsen CO, Jannasch HW (1991) Sulfide oxidation in the anoxic Black Sea chemocline. Deep-Sea Research, 38(2): 1083–1103

Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution, 30(4): 772–780

Kawaichi S, Ito N, Kamikawa R, Sugawara T, Yoshida T, Sako Y (2013) *Ardenticatena maritima* gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum '*Chloroflexi*' isolated from an iron-rich coastal hydrothermal field, and description of *Ardenticatenia* classis nov. International Jounral of Systematic and Evolutionary Microbiology, 63: 2992–3002

Kimber RL, Boothman C, Purdie P, Livens FR, Lloyd JR (2012) Biogeochemical behaviour of plutonium during anoxic biostimulation of contaminated sediments. Mineralogical Magazine, 76(3): 567–578

Kuever J (2014) The Family *Desulfobulbaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (eds) The Prokaryotes. Springer, Berlin, Heidelberg

Libes SM (1992) An Introduction to Marine Biogeochemistry. John Wiley and Sons, Hoboken, NJ: pp. 734

Liebgott P-P, Joseph M, Fardeau M-L, Cayol J-L, Falsen E, Chamkh F, Qatibi A-I, Labat M (2008) *Clostridiisalibacter paucivorans* gen. nov., sp. nov., a novel moderately halophilic bacterium isolated from olive mill wastewater. International Journal of Systematic and Evolutionary Microbiology, 58: 61–67

Loh PS, Reeves AD, Harvey SM, Overnell J, Miller AEJ (2008) The fate of terrestrial organic matter in two Scottish sea lochs. Estuarine, Coastal, and Shelf Science, 76: 566–579

Mahmood q, Zheng P, Cai J, Wu D, Hu B, Li J (2007) Anoxic sulfide biooxidation using nitrite as an electron acceptor. Journal of Hazardous Materials, 147: 239–256

Mohr KI, Garcia RO, Gerth K, Irschik H, Müller R (2012) *Sandaracinus amylolyticus* gen. nov., sp. nov., a start-degrading soil myxobacterium, and description of *Sandaracinaceae* fam. nov. International Journal of Systematic and Evolutionary Microbiology, 62: 1191–1198

Morford JL, Emerson S (1999) The geochemistry of redox sensitive trace metals in sediments. Geochimica et Cosmochimica Acta, 63(11-12): 1735-1750

Morton JT, Toran L, Edlund A, Metcalf JL, Lauber C, Knight R (2017) Uncovering the Horseshoe Effect in Microbial Analyses. mSystems, 2: https://doi.org/ 10.1128/mSystems.00166-16.

Morton JT, Sanders J, Quinn RA, McDonald D, Gonzalez A, Vázquez-Baeza Y, Navas-Molina JA, Son SJ, Metcalf JL, Hyde ER, Lladser M, Dorrestein PC, Knight R (2017) Balance Trees Reveal Microbial Niche Differentiation. mSystems, 2:e00162-16. https://doi.org/10.1128/mSystems.00162-16.

Nobu MK, Dodsworth JA, Murugapiran SK, Rinke C, Gies EA, Webster G, Schwientek P, Kille P, John Parkes R, Sass H, Jørgensen BB, Weightman AJ, Liu W-T, Hallam SJ, Tsiamis G, Woyke T, Hedlund BP (2016) Phylogeny and physiology of candidate phylum 'Atribacter' (OP9/JS1) inferred from cultivation-independent genomics. The ISME Journal, 10: 273–286

O'Reilly S, Szpak M, Monteys X, Flanagan P, Allen C, Kelleher B (2014) Hydrographic controls on marine organic matter fate and microbial diversity in the western Irish Sea. Geophysical Research Abstracts, 16.

Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE 5(3): e9490. https://doi.org/10.1371/journal.pone.0009490

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41(D1): 590–596

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ravot G, Magot M, Fardeau M-L, Patel BKC, Thomas P, Garcia J-L, Olliver B (1999) *Fusibacter paucivorans* gen. nov., sp. nov., an anaerobic, thiosulfate-reducing bacterium from an oil-producing well. International Journal of Systematic Bacteriology, 49: 1141–1147

Ridgway IM, Price NB (1987) Geochemical associations and post-depositional mobility of heavy metals in coastal sediments: Loch Etive, Scotland. Marine Chemistry, 21: 229–248

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584 https://doi.org/10.7717/peerj.2584

Roh Y, Chon C-M, Moon J-W (2007) Metal reduction and biomineralization by an alkaliphilic metal-reducing bacterium, *Alkaliphilus metalliredigens* (QYMF). Geosciences Journal, 11(4): 415–423

Rusin PA, Quintana L, Brainard JR, Strietelmeier BA, Talt CD, Ekberg SC, Palmer PD, Newton TW, Clark DL (1994) Solubilization of Plutonium Hydrous Oxide by Iron-Reducing Bacteria. Environmental Science and Technology, 28: 1686–1690

Serrano AE, Escudero LV, Tebes-Cayo C, Acosta M, Encalada O, Fernández-Moroso S, Demergasso C (2017) First draft genome sequence of a strain from the genus *Fusibacter* isolated from Salar de Ascotán in Northern Chile. Standards in Genomic Sciences, 12: https://doi.org/10.1186/s40793-017-0252-4

Slobodkin AI, Tourova TP, Kostrikina NA, Chernyh NA, Bonch-Osmolovskaya EA, Jeanthon C, Jones BE (2003) *Tepidibacter thalassicus* gen. nov., sp. nov., a novel moderately thermophilic, anaerobic, fermentative bacterium from a deep-sea hydrothermal vent. International Journal of Systematic and Evolutionary Microbiology, 53: 1131–1134

Smii L, Ben Hania W, Cayol J-L, Joseph M, Hamdi M, Olliver B, Fardeau M-L (2015) *Fusbacter bizertensis* sp. nov., isolated from a corroded kerosene storage tank. International Journal of Systematic and Evolutionary Microbiology, 65: 117–121

Suzuku D, Li Z, Cui X, Zhang C, Katayama A (2014) Reclassification of *Desulfobacterium anilini* as *Desulfatiglans anilini* comb. nov. within *Desulfatiglans* gen. nov., and description of a 4-chlorophenol-degrading sulfate-reducing bacterium, *Desulfatiglans parachlorophenolica* sp. nov.. International Journal of Systematic and Evolutionary Microbiology, 64: 3081–3086

Tan H-Q, Wu X-Y, Zhang X-Q, Wu M, Zhu Z-F (2012) *Tepidibacter mesophilus* sp. nov., a mesophilic fermentative anaerobe isolated from soil polluted by crude oil, and emended description of the genus *Tepidibacter*. International Journal of Systematic and Evolutionary Microbiology, 62: 66–70

Thamdrup B, Finster K, Fossing H, Würgler Hansen J, Barker Jørgensen B (1994) Thiosulfate and sulfite distirbutions in porewater of marine sediments related to manganese, iron, and sulfur geochemistry. Geochimica et Cosmochimica Acta, 58: 67–73

Thauer RK, Jungermann K, Decker K (1977) Energy Conservation in Chemotrophic Anaerobic Bacteria. Bacteriological Reviews, 41: 100–180. doi:10.1073/pnas.0803850105

Tuttle JH, Jannasch HW (1973) Sulfide and Thiosulfate-Oxidizing Bacteria in Anoxic Marine Basins. Marine Biology, 20: 64–70

Urios L, Cueff V, Pignet P, Barbier G (2004) *Tepidibacter formicigenes* sp. nov., a novel spore-forming bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent. International Journal of Systematic and Evolutionary Microbiology, 54: 439–443

Wang W, Vollertsen J, Hvitved-Jacobsen T (2005) Anoxic sulfide oxidation in wastewater of sewer networks. Water Science and Technology, 52(3): 191–199

Wood BJB, Tett PB, Edwards A (1973) An Introduction to the Phytoplankton, Primary Production and Relevant Hydrography of Loch Etive. Journal of Ecoloy, 61(2): 569–585

Wu X-Y, Shi K-L, Xu X-W, Wu M, Oren A, Zhu X-F (2010) *Alkaliphilus halophilus* sp. nov., a strictly anaerobic and halophilic bacterium isolated from a saline lake, and emended description of the genus *Alkaliphilus*. International Journal of Systematic and Evolutionary Microbiology, 60: 2892–2902

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y (2006) *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacgterial phylum *Chloroflexi*. International Journal of Systematic and Evolutionary Microbiology, 56: 1331–1340

Yang S-J, Choo Y-J, Cho J-C (2007) *Lutimonas vermicola* gen. nov., sp nov., a member of the family *Flavobacateriaceae* isolated from the marine polychaete *Periserulla leucophryna*. International Journal of Systematic and Evolutionary Microbiology, 57: 1679–1684

Yao W, Millero FJ (1995) Oxidation of hydrogen sulfide by Mn(IV) and Fe(III) (hydr)oxides in seawater. ACS Symposium Series, 612: 260–279

Yai W, Millero FJ (1996) Oxidation of hydrogen sulfide by hydrous Fe(III) oxides in seawater. Marine Chemistry, 52: 1–16

Zakharyuk A, Kozyreva L, Ariskina E, Troshina O, Kopitsyn D, Shcherbakova V (2017) *Alkaliphilus namsaraevii* sp. nov., an alkaliphilic iron- and sulfur-reducing bacterium isolated from a steppe soda lake. International Journal of Systematic and Evolutionary Microbiology, 67: 1990–1995

Zhilina TN, Zavarzina DG, Kolganova TV, Lysenko AM, Tourova TP (2009) *Alkaliphilus peptidofermentans* sp. nov., a New Alkaliphilic Bacterial Soda Lake Isolate Capable of Peptide Fermentation and Fe(III) Reduction. Microbiology, 78(4): 445–454

Chapter 6. Conclusions and Discussion for Future Work

6.1 General Conclusions Across All Depth Cores

In the previous chapters, the three sites were generally discussed in isolation, with little cross-references between different sites (except to provide context to certain index values). In this chapter, specific microbial taxonomic and functional, and geochemical analysis results will be compared across all the sites.

DNA-based methods of microbial community profiling enable a high-resolution (*e.g.*, on a 0.5 cm scale) and high-throughput (*e.g.*, millions of sequenced reads) (Makarova and Koonin, 2003; Rothberg and Leamon, 2008) analysis of indigenous microbial communities and can be used to link microbial community assemblage and function with inferred biogeochemical cycling in an environment (Zerkle *et al.*, 2005). This thesis aimed to infer and characterise the biogeochemical cycling of key nutrients and radionuclide contaminants of a number of depth cores collected from a range of environments, based on microbial community assemblage and functional profiles using DNA-based methods, and compare these inferences with directly measure geochemical analysis of the cores.

The hypothesis of this thesis was that microbial community structure, assemblage, and function would be vertically stratified based on the 'classical' model of redox stratification and diffusion of terminal electron acceptors (TEA), where microbial community assemblages would be controlled by the geochemical redox conditions of the surrounding environment, with TEA succession generally following the pattern of O_2, NO_x^-, Fe^{3+}, SO_4^{2-} reduction coupled to electron donor oxidation, with increasing depth (Thauer *et al.*, 1977; Berner, 1980; Conrad, 1996). Previous studies of similar cores have found very fine-scaled and short-spatial stratification of microbial community assemblages directly linked to the vertical diffusion of oxygen from the surface in mangrove soils (Lüdemann *et al.*, 1999), marine sediments (Durbin and Teske, 2011), geological strata (Lin *et al.*, 2011), and continental shelf sediments (Hunter *et al.*, 2005). Inferences about the biogeochemical cycling of nutrients, and of the transport and fate of radionuclide contaminants, would be based on the redox-linked structures of microbial communities as a function of depth. It was assumed *a priori* that thermodynamics would be the principal control of microbial community assemblages, *i.e.*, depth-related diffusion and availability of TEAs would be the most apparent control on microbial community structure, and profiles of these assemblages would in turn reflect changing redox conditions and TEA diffusion and

availability with depth. Vertical stratification of microbial community assemblages was observed in all cores, though not necessarily following the 'classical' model of O_2 through to SO_4^{2-} and methanogenesis, and the specific nature and extent of observed vertical stratifications were varied and based on several environmental factors unique to each site.

In order to accurately determine geochemical and redox conditions of a site using microbial taxonomic and functional profiles derived from *in situ* DNA-based analysis, strong but progressive spatial and successive changes in the patterns of community and functional structural changes must be observed. All cores from all sites resulted in the production of fine-scaled patterns of microbial community assemblage structural changes, as opposed to random or staggered distributions, enhancing the confidence that the observed microbial communities were constructs of longer-term geochemical, geophysical, and hydrological conditions, not a function of random amplification or sequencing. Consistent, progressive, fine-scale changes in microbial community structure profiles with increasing depth, that are reproducible across multiple methods, as observed in this thesis, indicate that microbial communities in the depth cores are structured and determined based on growth and selection for particular microbial assemblages, and reflect the stability of key environmental properties of a system. The additional benefit to using biomarkers to indicate geochemical conditions as opposed to measuring E_h directly is that these structured microbial communities, that represent growth and selection based on geochemical conditions, are indicative of a steady site in the medium-term, *i.e.*, weeks to months (Nielsen *et al.*, 2007), whereas directly measuring E_h can change on a very small spatial or temporal scale.

All microbial community profiles derived from DNA-based methods used at the sites resulted in community profiles that appeared to be structured differently and reflected environmental properties unique to each site. Furthermore, sites where certain environmental properties were less imposing, community profiles had less defined stratification and more homogeneity as a result of this lack of structure (*e.g.*, the Irish Sea Mud Patch, see Chapter 4). It became clear that microbial structures of the cores were impacted by not just diffusion and succession of TEAs, but, potentially more so, by different geophysical and hydrological settings unique to each core – or rather, that unique geophysical and hydrological properties at each site impacted not only diffusion of TEAs, but also electron donors and radionuclide contaminants.

Internal variations of microbial communities appeared to be explained and imposed by factors including highly-localised water movement, macrofaunal bioturbation, and the input (or lack of) carbon as an electron donor. The roles of hydrology and carbon inputs was especially pertinent at the Needle's Eye (Chapter 3) and Ravenglass (Chapter 4) sites.

The general trends in microbial community structures observed at each site reflected the geochemistry measured, as well as the location of each site, *e.g.*, the terrestrial, estuarine, and marine origins of each core. Taxonomic classifications of amplicon sequence variants (ASVs) were different between sites, and reflected microbial taxonomic assemblages expected based on the origin of the cores, including: the higher abundance of sulfate-reducing bacteria (SRB) in the marine depth cores than the terrestrial cores, as sulfate levels were tend to be higher in marine waters than in the freshwater (Capone and Kiene, 1988); and the higher abundance of methanogens in terrestrial environments than marine environments, likely due to the much higher TOC levels and; nitrogen cycling taxonomic and functional gene assemblages that were different between terrestrial and marine sites, *e.g.*, the high abundance of that bacterial nitrite oxidising taxa at Needle's Eye that were low in abundance at the marine sites (Stephen *et al.*, 1996). Common cosmopolitan marine-associated taxa dominated the marine depth cores, including the *Woeseia* spp. (Mußmann *et al.*, 2017) and numerous γ-*proteobacteria* (Edwards *et al.*, 2010), whereas common soil groups such as the *Acidobacteria* (Jones *et al.*, 2009) dominated the terrestrial Needle's Eye cores.

Further evidence the microbial community assemblages were being imposed by the surrounding geophysical and hydrological surroundings was presence or lack of complete catabolic cycles. The Needle's Eye contaminated core appeared to have complete nitrogen cycling, with the stratification of denitrification and nitrification at the surface, and DNRA and nitrogen-fixation at depth, followed by further nitrite reduction and nitrification at 35 cm. Furthermore, the emergence of methanogens between 24 – 34 cm in the Needle's Eye contaminated core precluded the emergence of anaerobic methane oxidisers at 35 cm. At Ravenglass, the two main regions of the core appeared to be completely separate and not linked, with 1 – 4 cm being characterised as having the 'classical' stratification of redox processes through to sulfate-reducing, and 10 – 28 cm characterised as being driven by hydrology, as mixed aerobic and anaerobic and becoming more aerobic with depth, likely reflecting

the highly mixed nature of the bottom of the core due to tidal influence. At the Mud Patch, evidence of mixing from bioturbation was found, and here, limited carbon availability did appear to have more of an impact on microbial community structure than hydrological features. Loch Etive represented the most complicated site in terms of interpreting the precise geochemical, geophysical, and hydrological characteristics of site from the microbial profiles. Both cores appeared to be uncoupled between the upper halves and lower halves of the cores, with the upper regions dominated from surface by sulfate reduction but demonstrating a transition to fermentation and sulfur oxidation at depth, with no immediately clear cause of this structure at depth.

At all the sites, microbial community structure showed gradual and progressive change as a function of depth, and taxonomic structures were directly linked to the presence of key biogeochemical cycling at each site, reflecting measured geochemical properties. For example, at Needle's Eye, Ravenglass, and Loch Etive Shallow Basin, porewater sulfate concentrations showed strong positive correlations to known sulfate-reducing δ-*proteobacteria* taxa, and at Ravenglass, the presence of *Ardenticatenales* taxa correlated to the reduction of solid phase iron, while at the Shallow Basin, the sulfur oxidising *Thiogranum* spp. positively correlated with thiosulfate concentrations. The presence of key microbial taxonomic groups with well-characterised function is therefore a reliable marker for the presence of the required electron donor and TEA of that taxonomic group.

6.2 The Benefits and Limitations of Using Microbial Functional Inference Tools

Another question of this thesis was the reliability of rudimentary molecular biological techniques at predicting the prevailing geochemical and redox conditions of a site by targeting a range of functional catabolic biogeochemical cycling genes with 'end point' PCR and comparing these to directly measured geochemical variables. This was taken a step further with the advent of Tax4Fun (Asshauer *et al.*, 2015), which creates predicted microbial functional from 16S rRNA gene sequencing taxonomic classifications. This enabled the targeting of not only the few genes targeted during 'end point' PCR, but a whole suite of genes including genes from whole catabolic pathways, to get a much more detailed analysis of functional gene profiles and make more specific and detailed predictions of redox conditions and biogeochemical cycling.

The question emerged of whether it is appropriate to use automated software tools such as Tax4Fun (or PICRUSt, Piphillin, and Paprica) that link taxonomic classifications to fully-sequenced genomes of the closest known relative and create an inferred functional profile of each ASV. The benefit that these tools provide is the automation of a previous step in 16S rRNA gene sequencing analysis which was identifying key functional microbial taxonomic monophyletic groups that performed known and specific respiratory and catabolic processes, to infer the presence of key microbial biogeochemical processes and reconstruct predicted redox processes. Previously, key microbial taxonomic classifications would be sought in the profiles, which in highly diverse and complex environments such as Needle's Eye, can be an exhaustive process. Then, representative sequences of key microbial indicator taxa would be entered into a BLAST search, and closely related matched sequences would be input into a phylogenetic tree in order to see the relatedness of sequences, as well as comparing the environmental conditions of the matched sequences, and literature searches of closely-related relatives. The inference of microbial function of uncultured sequences in 16S rRNA gene datasets from phylogenetically related cultured representatives is common practice (Zaneveld *et al.*, 2010; Segata and Huttenhower, 2010). Automating the inference of microbial functional capacity removes the need for many of these steps, as functional profiles can be assembled directly from 16S rRNA gene taxonomic profiles in a few seconds (Langille *et al.*, 2013; Asshauer *et al.*, 2015). Sets of genes and pathways could then be targeted automatically with a script that extracted abundances of target genes and assembled them into abundance depth profiles, as was the case in this thesis. Abundance profiles of key functional genes and pathways could then be compared to 'end point' PCR band intensities, as well as used directly to infer redox conditions and biogeochemical processes occurring in a site and compared to directly measured geochemical conditions. Furthermore, while taxonomic assemblages can differ between sites, the inferred function might not, *i.e.*, different taxonomic groups can perform the same function, *e.g.*, ammonia oxidation by bacteria and archaea. Creating functional profiles direct from taxonomy can be a way to identify key functional processes without having to exhaustively study subtle differences in microbial community taxonomic assemblages of sites.

The efficacy of this process was confirmed at Needle's Eye, where 'end point' PCRs, Tax4Fun inferred functional gene profiles, and geochemical conditions all

corroborated well between the two halves of the contaminated core. Furthermore, microbial functional profiles of the Needle's Eye site informed the prediction that uranium (U) contaminants in the 1 – 23 cm region were likely to be in the oxidised U^{6+} state. This seemed contradictory, as normally the accumulation of U would imply that it was likely in the reduced U^{4+} state, as U^{6+} is mobile and soluble and unlikely to accumulate while U^{4+} is immobile and insoluble (Lloyd, 2003). TOCs in the Needle's Eye depth core also lead to the supposition that the U at 8 cm would be reduced, as in a 'classical' vertical depth profile, such high TOCs would normally result in high microbial activity and the rapid turnover of TEAs from the surface downwards, including U (Lovley and Chapelle, 1995). However, all biological markers, including the Tax4Fun profiles, indicated that the U peak at 8 cm coincided with peak of oxidising taxa and genes, meaning that there was no evidence of any microbially-mediated U bioreduction of U. It was later confirmed by HERFD-XANES analysis that the U at 8 cm was indeed almost entirely oxidised U^{6+} (Chapter 3.3.6.3). The Needle's Eye contaminated core also demonstrated the high resolution of the functional inference methods, as at 35 cm, the spike in nitrogen cycling and methane cycling genes in the Tax4Fun profiles highlighted a dramatically different set of geochemical conditions limited almost exclusively to this depth increment. The result of this was the prediction that U reduced from 24 – 34 cm would be at least partly reoxidised at 35 cm, as was later confirmed by HERFD-XANES analysis. Although the exact nature and depths of the subsurface groundwater flow was not identified by the original BGS report (MacKenzie *et al.*, 1991), it is highly likely that the 8 and 35 cm increments, and spikes in aerobic processes, is driven by the introduction of oxic groundwaters at these depths. By comparison, the control core had much more progressive change in function with depth and lacked a clear step-change as observed at 23 – 24 cm in the contaminated core. This likely reflected different hydrological conditions at the site, *i.e.*, the contaminated core had more defined subsurface groundwater flow confined to small spatial zones such as 8 cm and 35 cm specifically, whereas the control core appeared to have evidence of a less-defined zone of groundwater flow.

At Ravenglass, the inferred functional gene profiles did not demonstrate a depth core controlled by vertical diffusion of oxygen and other TEAs throughout, but a core that between ~ 1 – 5 cm, was controlled by vertical diffusion, and from 10 – 28 cm, was a homogenous mix of functional processes. (Figures 4.17, 4.19, and 4.20).

This was complicated further by geochemical analysis which appeared to show mixed Mn, Fe, and sulfate reduction profiles that on their own would have resulted in an uncertain analysis of redox conditions (Figures 4.26 and 4.27). However, the high resolution functional gene profiles demonstrated that between 1 – 4 cm there was rapid succession to sulfate-reducing conditions at 4 cm, which can be seen in the solid phase Mn and Fe, and porewater sulfate levels (Figures 4.26 and 4.27), followed by a heterogenous zone characterised by the decrease in *dsrA* between 5 – 9 cm, evidenced by some Fe reoxidation (Figure 4.25), and lastly a homogenous 10 – 28 cm zone that is comprised of a mix of aerobic and anaerobic functional genes, with evidence of mixed Mn, Fe, and sulfate patterns. This 10 – 28 cm region is likely similar to the Needle's Eye in respect that it is likely a reflection of *in situ* hydrological conditions, in this case, from the tide, causing mixing of microbial communities and functions. In the case of Ravenglass, interpreting either of the Tax4Fun inferred gene or geochemical analysis profiles alone would have been difficult, without the context of the other. This is in contrast to Needle's Eye, where Tax4Fun profiles were able to clearly describe and predict geochemical characteristics of the site.

The Mud Patch was the least definitive of the sites and showed mixed functional profiles throughout. The only certainty that could be derived from Tax4Fun inferred gene profiles from the Mud Patch was the emergence of *dsrA* with depth, and the absence of methanogenic function (Figure 4.21). On their own, Tax4Fun profiles implied a depth core that was mixed aerobic and anaerobic function, and likely metabolically inactive due to limited carbon levels due to clearly defined stratification. In this case, geochemical analysis helped to clarify this analysis, by demonstrating only mild reduction of Mn, Fe, and sulfate (Figures 4.30 and 4.31).

The Loch Etive cores showed the most complicated profiles with depth, and while there was vertical stratification of functional gene assemblage, the cores did not show any sign of the 'classical' vertical stratification exhibited at Needle's Eye or in 1 – 4 cm or Ravenglass. Loch Etive represented the only site where Tax4Fun profiles directly contradicted observed geochemical properties, in that at depth in both cores, oxygen-dependent processes were inferred to increase (ammonia and methane oxidation genes) (Figures 5.18, Figure 5.20, Figure 5.21, Figure 5.23, Figure 5.26). This was not reflected in either core by the reoxidation of Fe or Mn (Figures 5.29, Figure 5.30, Figure 5.33, Figure 5.34). Instead, Loch Etive appeared to be characterised by anaerobic fermentation, but in the case of the Shallow Basin,

particularly by a complex sulfur cycling process that must involve the oxidation of reduced sulfur species (Chapter 5.4.2). Detailed analysis of this sulfur cycling was hampered by the fact that Tax4Fun has not been updated since 2015.

The ability of the Tax4Fun functional gene profiles to accurately predict geochemical and redox conditions of site mirrored the transition from terrestrial, to estuarine, to marine environments. Sites that had clearly-defined redox stratification was reflected in the Tax4Fun profiles, whereas sites with little stratification resulted in less-definitive Tax4Fun profiles that were more difficult to interpret accurately without geochemical context. In order for Tax4Fun profiles to be interpretable, the environmental system must have variation with depth that results in clear patterns of abundance of stratification with depth. However, in the case of Needle's Eye and Ravenglass, and to a lesser extent the Mud Patch, Tax4Fun functional gene profiles gave useful context to the geochemical analysis performed at that site and were able to give accurate predictions to the dominant biogeochemical processes occurring at the sites.

One clear outcome identified was that 16S rRNA gene sequencing targeting the V4 region on the Illumina platform resulted in better, more clearly-defined, inferred functional profiles than sequencing targeting the V4-V5 region on the Ion Torrent platform. Though the data are not presented in this thesis, the Needle's Eye depth core was originally sequenced using Ion Torrent, and later Illumina. The subsequent Tax4Fun profiles from the Illumina data gave much clearer functional gene profiles, particularly of the *amoA* gene, which better corroborated with the 'end point' PCR and geochemical profiles. At the Mud Patch and Loch Etive cores, where archaeal ammonia oxidisers would be expected to be present at surface as per the 'end point' PCR profiles, the *amoA* gene was inferred to have low and mixed abundances in the Tax4Fun profiles. This led to the hypothesis that Ion Torrent sequencing of the V4-V5 region of the 16S rRNA gene was leading to inferior archaeal sequencing and taxonomic classification that inhibited the ability of Tax4Fun to accurately infer archaeal *amoA* abundance in the depth cores, and that Illumina sequencing of the V4 region ameliorated this. The precise nature of this discrepancy between the sequencing methods is unclear, *i.e.*, whether it is a function of read length, where the longer reads of the V4-V5 region leads to inferior targeting of archaeal sequences, or if paired-end reading in Illumina resulted in better quality archaeal sequences and subsequently allowed for better taxonomic classification.

Regardless, it is important to note that many microbial studies have employed Ion Torrent sequencing, and these datasets will be in the literature (and continue to be published) for years to come, and this must be considered when functional inferences are made from these sequenced datasets.

One major issue with Tax4Fun is the constant and rapidly evolution of reference databases. Over time, databases will increase to include more resolved lineages where function is currently only inferred from distantly related taxonomic groups, more polyphyletic families with multiple catabolic and respiratory functions, and more currently unresolved lineages with no cultured representatives. Tax4Fun itself has not been updated since 2015 and relies on precomputed matrices built from SILVA123 and KEGG version 64, both of which have been updated several times since (to SILVA 132 and KEGG version 82 at time of writing). This means that many new taxonomic classifications and fully-sequenced genomes are missing. For example, the entire sulfur oxidising *sox* pathway is missing in Tax4Fun. It is hypothesised that sulfur oxidation is a key component of the Shallow Basin depth core, and it was not possible to create functional gene profiles of sulfur oxidation at this site due to Tax4Fun not being updated to include this pathway. As reference database and gene pathways continue to expand, the contextual interpretation by Tax4Fun will continue to become outdated and miss out on key functional pathways. PICRUSt has a similar shortcoming as it only uses the Greengenes 16S rRNA gene taxonomic reference database, which itself has not been updated since 2015. Other tools, such as Piphillin, have emerged that remove the need for precomputed matrices and can therefore simply be linked to the latest 16S rRNA taxonomic and KEGG reference databases when they are released. However, Piphillin is a closed-source software owned by a private company utilising their own computing platforms *via* the Internet, which may create issues for projects that are of a highly sensitive nature or that require the source code of the tool.

6.3 Using Taxonomic and Inferred Functional Profiles to Predict Radionuclide Redox Fate

Throughout this thesis, an overarching theme has been the ability to use microbial DNA as a biomarker to make inferences about the prevailing geochemical and redox conditions of the sites on the medium-term, and thus the anticipated longer-term redox transport and fate of a range of radionuclides, specifically

actinides. Specific predictions about actinide redox speciation in the sites based on the assemblage and stratification of key microbial taxonomic and phylogenetic groups were then corroborated either with direct speciation detection methods such as HERFD-XANES at Needle's Eye, or by measuring porewater and solid-phase actinide concentrations *via* ICP and XRF respectively such as at Ravenglass and Mud Patch.

At Needle's Eye, due to its clearly defined functional stratification with depth, even rudimentary techniques such as 'end point' PCR was able to construct functional profiles which appeared to conclusively show oxidising conditions between 0 – 20 cm, and very reducing conditions between 20 – 40 cm, and an oxidising incursion at ~ 35 cm, from which it was predicted that U would be predominantly 6+ between 0 – 20 cm, 4+ between 20 – 40 cm, and potentially reoxidised to 6+ at 35 cm. 16S rRNA gene sequencing profiles provided further evidence for these predictions. These predictions were validated by HERFD-XANES analysis.

At Ravenglass, 'end point' PCR and 16S rRNA gene profiles appeared to show rapid stratification of redox function between 0 – 6 cm, and it was therefore hypothesised that any Sellafield-derived actinide contaminants in this upper zone would be in the reduced tetravalent species. The lower 10 – 28 cm region of the core appeared to have homogenous distribution but heterogenous with regards to aerobic and anaerobic taxa and function. It was predicted that this region of the core was sub-oxic and not as reducing as ~ 4 cm, and therefore it was inconclusive as to the extent of any actinide reduction that may be occurring. However, geochemical analysis focused on Pu and Am, and therefore it was expected that Pu would be in the tetravalent species expected in sub-oxic conditions, and α-spectrometry showed the bulk of Pu/Am activity to be accumulated in the lower ~ 20 – 28 cm region of the core, and therefore likely to be in the insoluble and immobile 4+ species. At the Mud Patch, microbial profiles appeared to depict a core that was homogenous in distribution with depth, heterogenous with aerobic and sub-oxic processes, and with relatively low carbon and microbial activity (due to the lack of clearly defined stratification). Any actinides present here were predicted to be oxidised, or in the case of Pu, in the tetravalent species. This was corroborated with α-spectrometry. A lack of direct speciation measurement of Pu meant speciation could only be inferred by accumulation in the sediment being as a result of the Pu being immobile and insoluble, and therefore as Pu^{4+}.

With the Loch Etive Shallow Basin depth core, it was possible to make specific predictions about the redox speciation of actinides in the core based on microbial profiles. Taxonomic and functional profiles conclusively depicted a core that was initially sulfidic, becoming more reducing and fermenting with depth, resulting in anticipated reduction of metals, actinides, and sulfate. Any actinides present in this core were expected to be reduced either by direct enzymatic microbial bioreduction, or abiotically reduced by other microbially-reduced elements such as Mn/Fe^{2+} or reduced sulfur species. Al-Qasmi *et al*. (2018) performed a range of geochemical measurements and measured concentrations of a range of actinides such as U, Pu, and Am. The results presented in Al-Qasmi *et al*. corroborate with the results in this thesis that the core is reducing and that the actinides present are likely reduced, although the authors go as far as to predict that the Pu in the lower half of the core is reduced all the way to Pu^{3+} due to microbial sulfate reduction. This conflicts with the microbial profiles that show sulfate reduction confined to the upper 10 cm of the core. Without direct Pu speciation measurements, it is not possible to discern the accuracy of the prediction of Pu^{3+} in Al-Qasmi *et al*. At the Deep Basin, it was not possible to corroborate microbial profiles and the predictions of actinide speciation with the geochemical measurements in Al-Qasmi *et al*. This is somewhat unsurprising as the microbial profiles in the Deep Basin lacked clearly defined stratification of redox couples, and instead showed a heterogenous mix of sulfate reduction and fermentation, with evidence of oxidation at the bottom of the core.

Using microbial DNA as a biomarker to construct taxonomic and functional profiles and link the stratification of microbial phylogeny and function to the prevailing geochemical conditions of a site proved to be an effective method for predicting the redox transport and fate of radionuclides, specifically actinides, in the sites tested in this thesis. The efficacy of the method is heavily influenced by the nature of the site. Predictions of expected actinide speciation based on microbial profiles at sites with clearly defined functional stratification with depth, such as Needle's Eye, Ravenglass, and the Shallow Basin, proved to be highly accurate. Furthermore, at sites with gradual, fine-scaled microbial distributions with depth, these methods were able to detect changes in redox conditions at the cm scale, and therefore make predictions about changes in actinide speciation on the cm scale also – such as at Needle's Eye and the 35 cm incursion of oxidising conditions, likely pertaining to anaerobic methane oxidation and microaerophilic conditions induced by groundwater flow.

However, at sites with less clearly defined stratification with depth, sites with low metabolic activity, or sites that are heterogenous in profile due to physical mixing or bioturbation, it was more difficult to make precise predictions about actinide speciation based solely on microbial profiles. At these sites, microbial taxonomic and functional profiles can be used to confirm and corroborate with directly measured geochemical analyses. This proved to be the case at the Mud Patch, but at the Deep Basin, microbial profiles appeared to contradict inferences made from geochemical analyses in Al-Qasmi *et al*. (2018). The only way to reconcile the discrepancies at the Deep Basin site would be to perform more geochemical and microbial analyses, such as HERFD-XANES, and qPCR and metatranscriptomics respectively.

6.4 Future Work and Considerations

One limitation of this thesis was a more detailed comparison of 'end point' PCR, Tax4Fun, and shotgun metagenome sequencing functional gene profiles. A more comprehensive comparison, that included more samples sequenced *via* shotgun metagenome sequencing, and from more depth cores, would have enabled a more detailed analysis of the accuracy of 'end point' PCR and Tax4Fun profiles compared to directly measured shotgun metagenome functional gene profiles. Another limitation of the shotgun metagenome sequencing that was performed in this thesis was the sequencing depth. The Needle's Eye depth core was a highly diverse environment, containing many functional processes occurring concurrently, and shotgun metagenome sequencing was simply not performed to a sufficiently deep level to adequately sequence and create functional gene profiles that reflected that highly diverse function. For example, only 48 *mcrA* gene reads were identified in by GraftM at 31 cm, the peak of the anaerobic conditions, in the Needle's Eye contaminated core. At this same depth increment, 'end point' PCR and Tax4Fun resulted in the *mcrA* gene being highly abundant (Figures 3.4F and 3.18E). Other genes, such as the *amoA*, were completely missing. This indicates that much deeper metagenome sequencing is required in order to fairly compare with the other methods. However, as of writing, shotgun metagenome sequencing is still considerably more expensive than 16S rRNA gene sequencing, especially to an adequate depth required for a site like Needle's Eye. One alternative approach could be to sequence only *via* shotgun metagenome sequencing, which would also sequence 16S rRNA gene fragments, and use the constructed 16S rRNA gene

taxonomic profiles to input into Tax4Fun, in an attempt to broadly compensate for insufficient deep metagenome sequencing.

Another limitation of this thesis was the time dedicated to understanding the taxonomic profiles derived from fungal ITS sequencing. Although a method of fungal sequencing using updated primers that targeted the ITS2 region (Toju *et al.*, 2012) on the Ion Torrent platform and a subsequent QIIME2 pipeline, as well as comparing that sequencing to sequencing performed following the ITS1 EMP protocol on Illumina, was developed and performed as part of this thesis, it was not possible to invest the time to properly investigate the fungal communities sequenced in the Needle's Eye contaminated core. Furthermore, fungal reference databases remain severely under-represented compared to bacterial and archaeal reference databases, making accurate taxonomic classification much more difficult for fungal sequencing compared to bacterial and archaeal sequencing. Specific known catabolic functions of fungal are also less well understood compared to bacteria and archaea, probably due to culturing. Lastly, due to the hypervariability of the ITS gene, it is not currently possible to accurately align sequenced ITS reads and therefore not possible to perform phylogenetic assessment of sequences. A targeted approach that could focus on aligning more closely related sequences, *e.g.*, at phylum or class level, where there is less hypervariability, could be explored.

6.4 References

Al-Qasmi H, Law GTW, Fifield LK, Howe JA, Brand T, Cowie GL, Law KA, Livens FR (2018) Deposition of artificial radionuclides in sediments of Loch Etive, Scotland. Journal of Environmental Radioactivity, 187: 45–52

Capone DG, Kiene RP (1988) Comparison of microbial dynamics in marine and freshwater sediments: Contrasts in anaerobic carbon catabolism. Limnology and Oceanography, 33(4): 725–749

Conrad R (1996) Soil Microorganisms as Controllers of Atmospheric Trace Gases (H2, CO, CH4, OCS, N2O, and NO). Microbiological Reviews, 60(4): 609–640

Durbin AM, Teske A (2011) Microbial diversity and stratification of South Pacific abyssal marine sediments. Environmental Microbiology, 13(12): 3219–3234

Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. PNAS, 102(41) 14683–14688

Hunter EM, Mills HJ, Kostka JE (2006) Microbial Community Diversity Associated with Carbon and Nitrogen Cycling in Permeable Shelf Sediments. Applied and Environmental Microbiology, 72(9): 5689–5701

Jones RT, Robeson MS, Lauber CL, Hamady M, Knight R, Fierer N (2009) A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. The ISME Journal, 3: 442–453

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C (2013) Predictive functional profiling of microbial communities using 16s rRNA marker gene sequences. Nature Biotechnology, 31(9): 814–823

Lin X, Kennedy D, Fredrickson J, Bjornstad B, Konopka A (2012) Vertical stratification of subsurface microbial community composition across geological formations at the Hanford Site. Environmental Microbiology, 14(2): 414–425

Lovley DR, Chapelle FH (1995) Deep subsurface microbial processes. Reviews in Geophysics, 33(3): 365–381

Lüdemann H, Arth I, Liesack W (2000) Spatial Changes in the Bacterial Community Structure along a Vertical Oxygen Gradient in Flooded Paddy Soil Cores. Applied and Environmental Microbiology, 66(2): 754–762

Makarova KS, Koonin EV (2003) Comparative genomics of archaea: how much have we learned in six years, and what's next? Genome Biology, 4(8): 115ussm

Mussmann M, Pjevac P, Krüger K, Dyksma S (2017) Genomic repertoire of the *Woeseiaceae*/JTB255 cosmopolitan and abundant core members of microbial communities in marine sediments. The ISME Journal, 11: 1276–1281

Nielsen KM, Johnsen PJ, Bensasson D, Daffonchio D (2007) Release and persistence of extracellular DNA in the environment. Environmental Biosafety Research, 6: 37–53

Rothberg JM, Leamon JH (2008) The development and impact of 454 sequencing. Nature Biotechnology, 26(10): 1117–1125

Segata N, Huttenhower C (2010) Toward an Efficient Method of Identifying Core Genes for Evolutionary and Functional Microbial Phylogenies. PLoS ONE, 6(9): e24704. doi:10.1371/journal.pone.0024704

Stephen JR, McCaig AE, Smith Z, Prosser JI, Embley TM (1996) Molecular Diversity of Soil and Marine 16S rRNA Gene Sequences Related to β-Subgroup Ammonia-Oxidizing Bacteria. Applied and Environmental Microbiology, 62(11): 4147–4154

Thauer RK, Jungermann K, Decker K (1977) Energy Conservation in Chemotrophic Anaerobic Bacteria. Bacteriological Reviews, 41: 100–180. doi:10.1073/pnas.0803850105

Zaneveld JR, Lozupone C, Gordon JI, Knight R (2010) Ribosomal RNA diversity predicts genome diversity in gut bacteria and their relatives. Nucleic Adics Reseach, 38(12): 3869–3879

Zerkle AL, House CH, Brantley SL (2005) Biogeochemical signatures through time as inferred from whole microbial genomes. American Journal of Science, 305: 467–502

Appendix A. Needle's Eye: AOV and Tukey HSD Tables

Appendix A.1. Contaminated Core: Summary of *p* values from one-way analysis (AOV) of variance and post hoc Tukey HSD of extracted DNA volumes across and between sample groups from: (A). Volumes of DNA extracted from each sample, and; (B). TOC volumes measured in each sample. Dash denotes p values between 0.05 – 0.1, * denotes 0.01 – 0.05, ** denotes 0.001 – 0.01, and *** denotes < 0.001.

A. DNA (ng/μL) AOV p-value = < 0.001 ***				**B. TOC (%) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.885			**2**	0.001 **		
3	< 0.001 ***	< 0.001 ***		**3**	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.519	**4**	< 0.001 ***	< 0.001 ***	0.187

Appendix A.2. Control Core: Summary of *p* values from one-way analysis (AOV) of variance and post hoc Tukey HSD of extracted DNA volumes across and between sample groups from volumes of DNA extracted from each sample.

DNA (ng/μL) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	< 0.001 ***		
3	< 0.001 ***	0.003 **	
4	< 0.001 ***	0.004 **	1

Appendix A.3. Contaminated Core: Summary of *p* values from AOV and Tukey HSD on 'end point' PCR functional gene band intensities as calculated in BioNumerics, across and between sample groups.

A. Bacterial ammonia monooxygenase – *amoA* AOV p-value = < 0.001 ***				**B. Archaeal ammonia monooxygenase – *amoA* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.127		
3	< 0.001 ***	0.554		**3**	< 0.001 ***	0.009 **	
4	< 0.001 ***	0.979	0.315	**4**	< 0.001 ***	< 0.001 ***	0.337
C. Nitrite reductase – *nirS* AOV p-value = < 0.001 ***				**D. Nitrite reductase – *nirK* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	1			**2**	0.215		
3	0.116	0.138		**3**	0.002 **	0.241	
4	< 0.001 ***	< 0.001 ***	0.003 **	**4**	0.008 **	0.506	0.945
E. Dissimilatory sulfite reductase – *dsrA* AOV p-value = < 0.001 ***				**F. Methyl Coenzyme M reductase – *mcrA* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	1			**2**	0.205		
3	< 0.001 ***	< 0.001 ***		**3**	0.008 **	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.861	**4**	0.006 **	< 0.001 ***	1

Appendix A.4. Control Core: Summary of *p* values from AOV and Tukey HSD on 'end point' PCR functional gene band intensities as calculated in BioNumerics, across and between sample groups.

A. Bacterial ammonia monooxygenase – *amoA* AOV p-value = < 0.001 ***				**B. Archaeal ammonia monooxygenase AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	< 0.001 ***		
3	< 0.001 ***	1		**3**	< 0.001 ***	0.857	
4	< 0.001 ***	0.252	0.281	**4**	< 0.001 ***	0.255	0.712
C. Nitrite reductase – *nirS* AOV p-value = < 0.001 ***				**D. Nitrite reductase – *nirK* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	1		
3	< 0.001 ***	0.065 -		**3**	0.142	0.174	
4	< 0.001 ***	0.012 *	0.920	**4**	< 0.001 ***	< 0.001 ***	0.001 **
E. Dissimilatory sulfite reductase – *dsrA* AOV p-value = < 0.001 ***				**F. Methyl Coenzyme M reductase – *mcrA* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.001 **			**2**	< 0.001 ***		
3	< 0.001 ***	0.462		**3**	0.064 -	0.023 *	
4	0.237	0.127	0.003 **	**4**	0.109	0.009 **	0.991

Appendix A.5. Summary of p-values derived from AOV of three alpha diversity indices between all respective sample groups within each core, and Tukey HSD between sample groups of each core.

Contaminated Core				**Control Core**			
A. Shannon Species Richness AOV p-value = < 0.001 ***				**B. Shannon Species Richness AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.003 **		
3	< 0.001 ***	0.562		**3**	0.022 *	0.854	
4	< 0.001 ***	0.030 *	0.403	**4**	< 0.001 ***	0.004 **	< 0.001 ***
C. Faith Phylogenetic Diversity AOV p-value = 0.063 -				**D. Faith Phylogenetic Diversity AOV p-value = 0.071 -**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.074 -			**2**	1		
3	0.106	0.998		**3**	0.597	0.584	
4	0.251	0.895	0.952	**4**	0.461	0.519	0.043 *
E. Pielou Species Evenness AOV p-value = < 0.001 ***				**F. Pielou Species Evenness AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.041 *			**2**	< 0.001 ***		
3	0.004 **	0.788		**3**	< 0.001 ***	0.122	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***	**4**	< 0.001 ***	< 0.001 ***	< 0.001 ***

Appendix A.6. Contaminated Core: Summary of *p* values derived from AOV and *post hoc* Tukey HSD of ASV abundance collapsed at genus level across and between the four depth sample groups.

A. Uncultured *Acidobacteria* Subgroup 6
AOV p-value = < 0.001 *

Group	1	2	3
2	0.299		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.039 *

B. Uncultured *Anaerolineaceae*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.206	
4	0.064 -	0.008 **	< 0.001 ***

C. Uncultured *Chloroflexi* KD4-96
AOV p-value = < 0.001 *

Group	1	2	3
2	0.966		
3	0.002 **	0.003 **	
4	< 0.001 ***	< 0.001 ***	0.283

D. *Sulfurifustis* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	0.004 **		
3	0.152	0.407	
4	0.888	< 0.001 ***	0.023 *

E. Uncultured methanogenic *Bathyarchaeia*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.035 *		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	0.007 **	0.477

F. Uncultured δ-*proteobacteria* SVA0485
AOV p-value = < 0.001 *

Group	1	2	3
2	0.987		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.002 **

G. Uncultured *Rokubacteria* NC10
AOV p-value = < 0.001 *

Group	1	2	3
2	1		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	1

H. Uncultured δ-*proteobacteria* MBNT15
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.120	
4	0.002 **	0.724	0.012 *

I. Uncultured *Bathyarchaeia Thermoprotei*
AOV p-value = < 0.001 *

Group	1	2	3
2	1		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.016 *

J. Uncultured *Thermodesulfovibriona*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.545		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.155

K. Uncultured *Nitrosopumilaceae*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.001 **		
3	0.051 -	< 0.001 ***	
4	0.001 **	< 0.001 ***	0.486

L. Candidatus *Methylomirabilis* spp.
AOV p-value = 0.556

Group	1	2	3
2	0.767		
3	1	0.765	
4	0.707	1	0.703

M. Uncultured *Nitrosomonadaceae* MND1
AOV p-value = < 0.001 *

Group	1	2	3
2	0.054 -		
3	0.014 *	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.064 -

N. *Methanosaeta* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	0.943		
3	0.018 *	0.003 **	
4	< 0.001 ***	< 0.001 ***	0.003 **

O. *Methanolinea* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	0.911		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

P. *Methanoregula* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	0.178		
3	0.036 *	< 0.001 ***	
4	0.015 *	< 0.001 ***	0.992

Appendix A.7. Control Core: Summary of *p* values derived from AOV and *post hoc* Tukey HSD of ASV abundance collapsed at genus level across and between the four depth sample groups.

A. Uncultured *Acidobacteria* Subgroup 6
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.220	
4	< 0.001 ***	0.007 **	0.467

B. Uncultured methanogenic *Bathyarchaeia*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.007 **		
3	< 0.001 ***	0.682	
4	< 0.001 ***	0.070 -	0.510

C. Uncultured *Gallionellaceae*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.001 **		
3	< 0.001 ***	0.034 *	
4	< 0.001 ***	< 0.001 ***	0.005 **

D. Uncultured α-*proteobacteria*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.767		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

E. Uncultured *Anaerolineaceae*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.048 *	
4	< 0.001 ***	0.974	0.106

F. Uncultured *Nitrospirae* 4-29-1
AOV p-value = < 0.001 *

Group	1	2	3
2	0.399		
3	0.011 *	0.336	
4	< 0.001 ***	< 0.001 ***	0.007 **

G. Uncultured *Acidobacteriales*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.008 **

H. Uncultured δ-*proteobacteria* MBNT15
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	0.019 *	0.049 *	
4	0.062 -	0.011 *	0.944

I. Uncultured *Rokubacteriales* NC10
AOV p-value = < 0.001 *

Group	1	2	3
2	0.004 **		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	0.006 **	0.736

J. Uncultured *Thermodesulfovibriona*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.235	
4	< 0.001 ***	< 0.001 ***	0.020 *

K. *Desulfobacca* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.881

L. *Pseudolabrys* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.147	
4	< 0.001 ***	< 0.001 ***	0.107

M. *Sideroxydans* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	0.994		
3	0.614	0.458	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

N. Uncultured *Micropepsaceae*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.003 **	
4	< 0.001 ***	0.072 -	0.532

O. *Nitrospira* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	0.001 **		
3	0.976	0.004 **	
4	0.398	0.063 -	0.653

P. Candidatus *Nitrosotalea* spp.
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.999	
4	0.001 **	0.989	0.999

Appendix A.8. Contaminated Core: Summary of p values from AOV and post hoc Tukey HSD on Tax4Fun inferred gene abundances of the genes targeted by 'end point' PCR across and between the four sample resolution groups.

A. Ammonia monooxygenase – *amoA*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.983		
3	0.009 **	0.134 *	
4	< 0.001 ***	< 0.001 ***	0.272

B. Nitrite reductase – *nirS*
AOV p-value = 0.287

Group	1	2	3
2	0.315		
3	0.373	1	
4	0.826	0.781	0.833

C. Nitrite reductase – *nirK*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.988		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.007 **

D. Dissimilatory sulfite reductase – *dsrA*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.930		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

E. Methyl Coenzyme M reductase – *mcrA*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.776		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.960

Appendix A.9. Contaminated Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of nitrogen metabolism genes across and between the four sample resolution groups.

A. Nitrate reductase – *narG*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.929		
3	0.017 *	0.002 **	
4	< 0.001 ***	< 0.001 ***	0.483

B. Nitrate reductase – *napA*
AOV p-value = 0.766

Group	1	2	3
2	0.928		
3	0.903	1	
4	0.714	0.963	0.982

C. Nitric oxide reductase – *norB*
AOV p-value = 0.959

Group	1	2	3
2	0.996		
3	1	0.990	
4	0.976	0.997	0.961

D. Nitrous-oxide reductase – *nosZ*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.916		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.928

E. Nitrite reductase – *nrfA*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.988		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

F. Nitrogenase – *nifD*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.999		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.009 **

Appendix A.10. Contaminated Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of genes across and between the four sample resolution groups.

A. Acetyl-CoA synthase – *acsB*
AOV p-value = < 0.001 *

Group	1	2	3
2	1		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

B. Adenylylsulfate reductase – *aprA*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.877		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

C. Methane monooxygenase – *mmoX*
AOV p-value = 0.006 **

Group	1	2	3
2	0.998		
3	0.556	0.629	
4	0.012 *	0.012 *	0.203

D. Methanol dehydrogenase – *mxaF*
AOV p-value = 0.660

Group	1	2	3
2	0.924		
3	0.929	1	
4	0.591	0.904	0.910

E. NiFe hydrogenase – *hydB*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.503		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.032 *

F. Methylene-tetrahydromethanopterin dehydrogenase – *mtdB*
AOV p-value = 0.905

Group	1	2	3
2	0.903		
3	0.983	0.989	
4	0.926	1	0.995

G. Heterodisulfide reductase – *hdrA*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.990		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

H. Anaerobic carbon-monoxide dehydrogenase – *cooS*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.803		
3	0.002 **	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.042 *

Appendix A.11. Control Core: Summary of p values from AOV and post hoc Tukey HSD on Tax4Fun inferred gene abundances of the genes targeted by 'end point' PCR across and between the four sample resolution groups.

A. Ammonia monooxygenase – *amoA*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	0.998	< 0.001 ***	
4	0.002 **	< 0.001 ***	0.001 **

B. Nitrite reductase – *nirS*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.598	
4	0.154	< 0.001 ***	0.019 *

C. Nitrite reductase – *nirK*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.793	
4	< 0.001 ***	0.459	0.949

D. Dissimilatory sulfite reductase – *dsrA*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.249	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

E. Methyl Coenzyme M reductase – *mcrA*
AOV p-value = 0.017 *

Group	1	2	3
2	0.026 *		
3	0.921	0.110	
4	0.107	0.902	0.340

Appendix A.12. Control Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of nitrogen metabolism genes across and between the four sample resolution groups.

A. Nitrate reductase – *narG* AOV p-value = < 0.001 ***				**B. Nitrate reductase – *napA* AOV p-value = 0.321**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.647		
3	< 0.001 ***	0.581		**3**	1	0.706	
4	< 0.001 ***	0.394	0.992	**4**	0.902	0.252	0.861
C. Nitric oxide reductase – *norB* AOV p-value = 0.013 *				**D. Nitrous-oxide reductase – *nosZ* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.022 *			**2**	< 0.001 ***		
3	0.060 -	0.975		**3**	< 0.001 ***	1	
4	0.797	0.147	0.310	**4**	< 0.001 ***	0.982	0.978
E. Nitrite reductase – *nrfA* AOV p-value = < 0.001 ***				**F. Nitrogenase – *nifD* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	< 0.001 ***		
3	< 0.001 ***	1		**3**	< 0.001 ***	0.267	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***	**4**	< 0.001 ***	< 0.001 ***	< 0.001 ***

Appendix A.13. Control Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of genes across and between the four sample resolution groups.

A. Acetyl-CoA synthase – *acsB* AOV p-value = < 0.001 ***				**B. Adenylylsulfate reductase – *aprA* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	< 0.001 ***		
3	< 0.001 ***	0.149		**3**	< 0.001 ***	0.410	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***	**4**	< 0.001 ***	< 0.001 ***	< 0.001 ***
C. Methane monooxygenase – *mmoX* AOV p-value = < 0.001 ***				**D. Methanol dehydrogenase – *mxaF* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.622		
3	< 0.001 ***	0.291		**3**	0.514	0.061 -	
4	< 0.001 ***	0.460	0.984	**4**	< 0.001 ***	< 0.001 ***	0.018 *
E. NiFe hydrogenase – *hydB* AOV p-value = < 0.001 ***				**F. Methylene-tetrahydromethanopterin dehydrogenase – *mtdB* AOV p-value = 0.003 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.008 **		
3	< 0.001 ***	< 0.001 ***		**3**	0.469	0.214	
4	< 0.001 ***	< 0.001 ***	0.011 *	**4**	1	0.006 **	0.430
G. Heterodisulfide reductase – *hdrA* AOV p-value = < 0.001 ***				**H. Anaerobic carbon-monoxide dehydrogenase – *cooS* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	< 0.001 ***		
3	< 0.001 ***	0.193		**3**	< 0.001 ***	0.004 **	
4	< 0.001 ***	< 0.001 ***	0.001 **	**4**	< 0.001 ***	< 0.001 ***	0.025 *

Appendix A.14. Contaminated Core: Summary of p values from AOV and post hoc Tukey HSD of porewater geochemical data across and between the four sample resolution groups.

A. Uranium
AOV p-value = < 0.001 *

Group	1	2	3
2	0.981		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.794

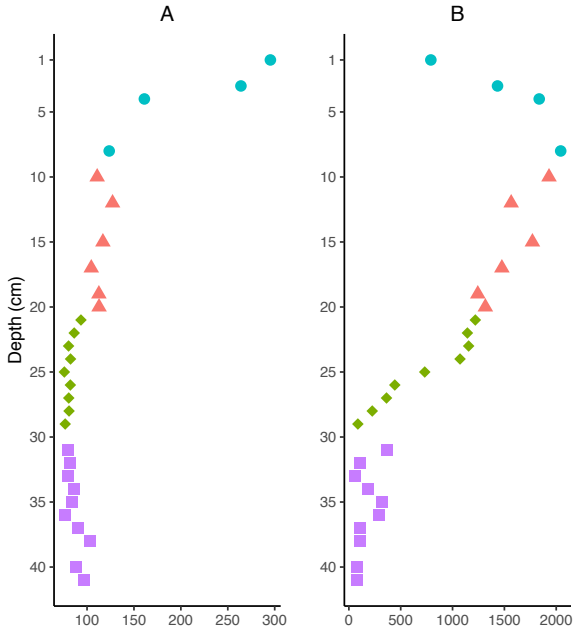
B. Manganese
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.227	
4	< 0.001 ***	0.343	0.978

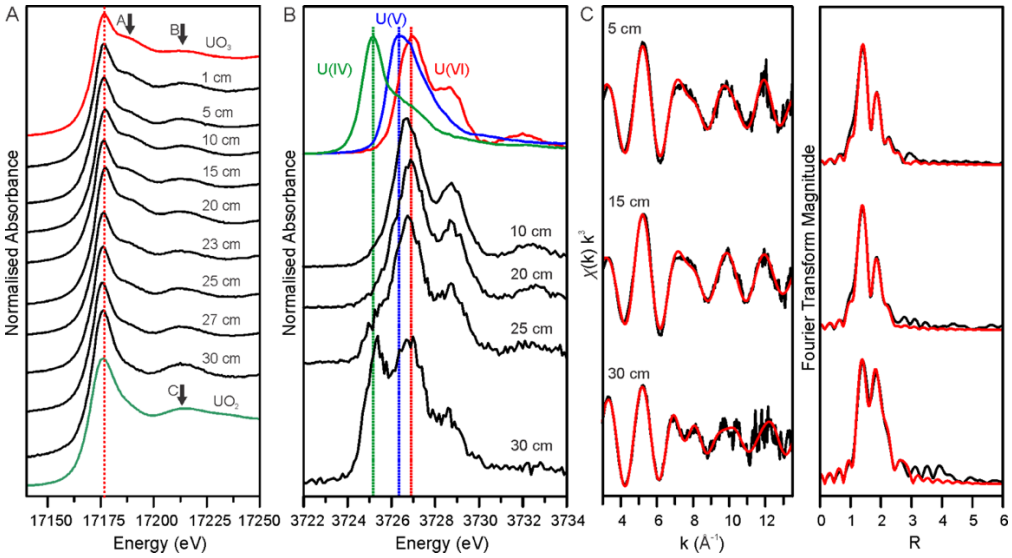
C. Zinc
AOV p-value = < 0.001 *

Group	1	2	3
2	0.954		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.011 *

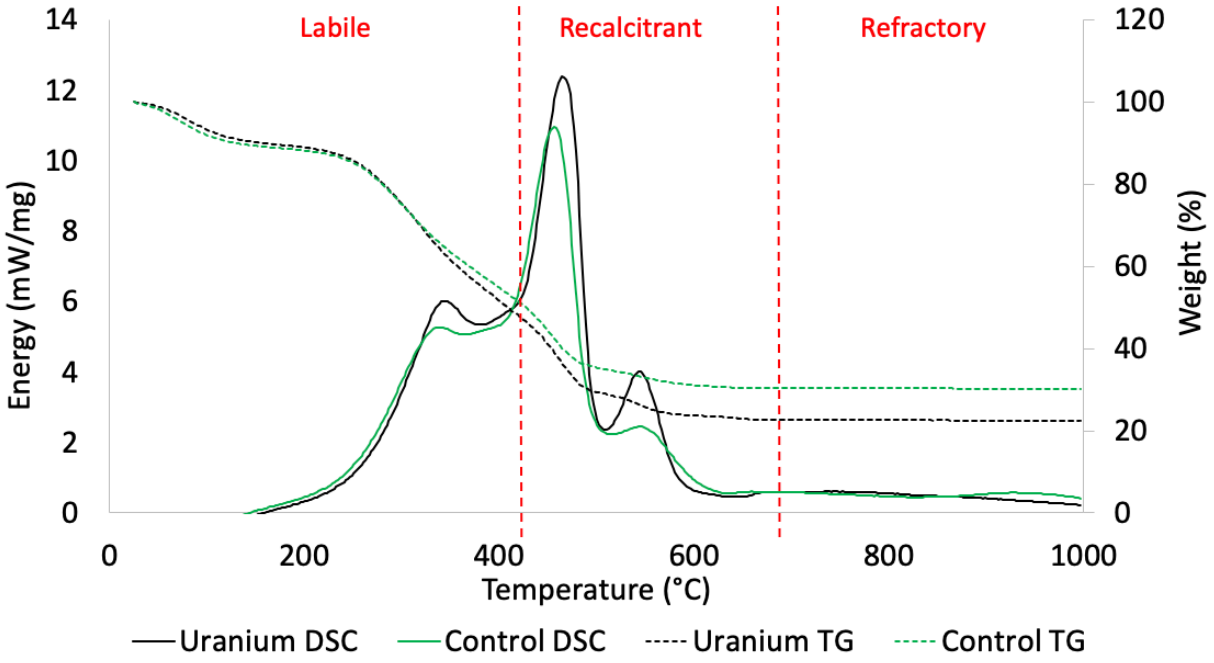
Appendix B. Needle's Eye: Geochemical Plots



Appendix B.1. Contaminated Core: Depth profiles of porewater: (A). Mn concentration as a function of depth, and (B). Zn concentrations. Figures are presented as ppm.

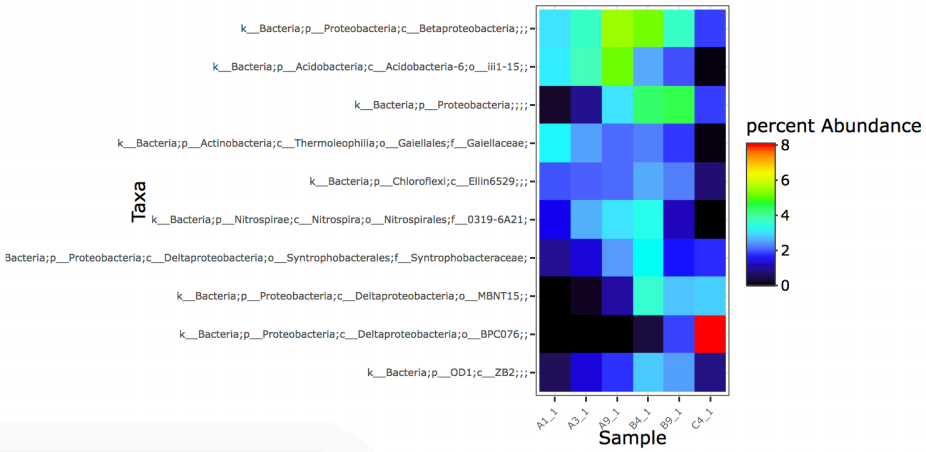


Appendix B.2. (A). U L3-edge XANES spectra of samples from different depths. UO3 and UO_2 are provided as standards for U^{6+} and U^{4+} respectively. For clarity the edge energy of U^{6+} is marked by a dotted line and the U^{6+} and U^{4+} XANES distinctive structural features are labelled with arrows. A = axial oxygen shoulder feature, B = equatorial oxygen feature of U^{6+}, C = equatorial oxygen feature of U^{4+} showing increased amplitude compared to B. (B). U M_4-edge HERFD-XANES from selected soil depths. Standards of U^{4+} (green, UO_2), $U^{5+/6+}$ (U_3O_8), and U^{6+} (red, UO_3) are provided for comparison. Dotted lines mark the edge energies of U in its various oxidation states where green = U^{4+}, blue = U^{5+}, and red = U^{6+}. (C). U L_3 edge EXAFS (black) with accompanying best-fits (red) for samples taken at 5, 15 and 30 cm depth. Data presented as k^3 with accompanying non-phase-shifted Fourier transform.

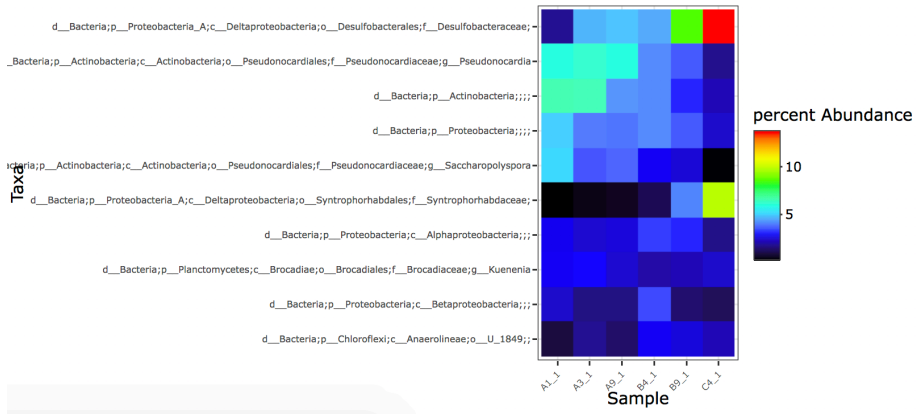


Appendix B.3. Thermograviometric and differential scanning calorimetry (TG-DSC) plot. Black solid line denotes contaminated core DSC, and black dotted line contaminated core TG. Green solid line denotes control core DSC, and green dotted line control core TG. Right y axis, Weight (%), refers to TG analysis. Energy (mW/mg), refers to DSC analysis. Red dotted lines refer to boundary temperature ranged which could be attributed to the labile, recalcitrant, and refractory organic carbon components of the samples in the TG mass loss curves.

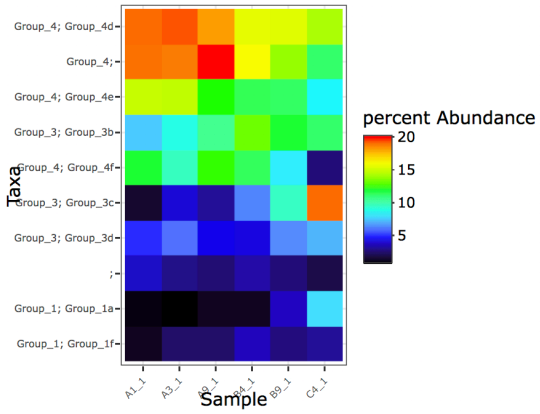
Appendix C. Needle's Eye: GraftM Heatmaps



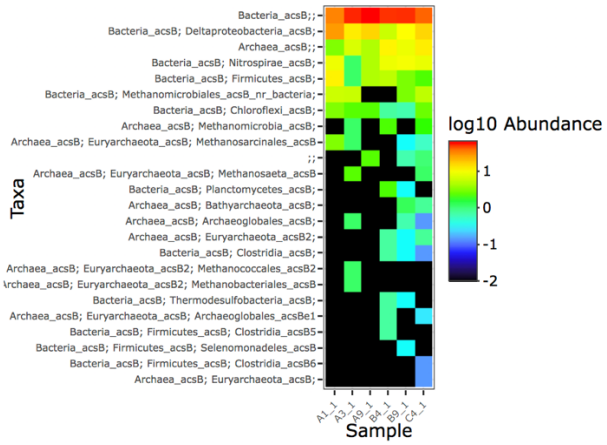
Appendix C.1. Heatmap of the 10 highest relative abundances of 16S rRNA gene taxonomies collapsed to genus level from shotgun metagenome sequencing as detected by GraftM. Sample A1_1 = 1 cm; A3_1 = 3 cm; A9_1 = 9 cm; B4_1 = 17 cm; B9_1 = 23 cm; C4_1 = 31 cm.



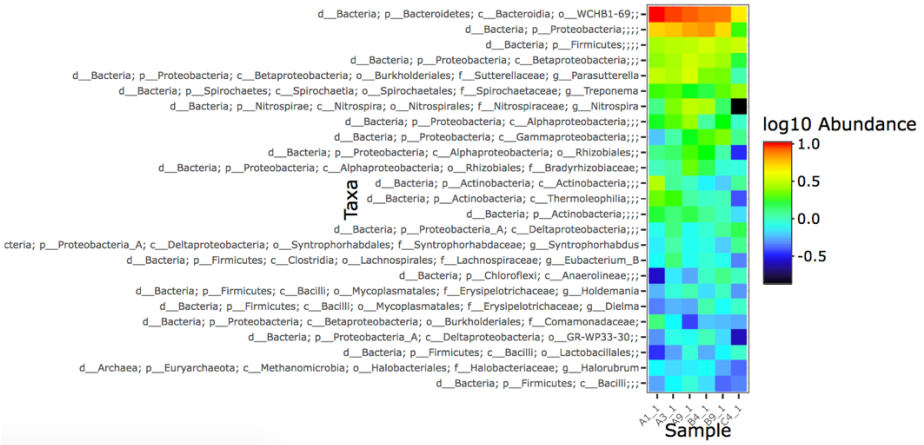
Appendix C.2. Heatmap of the 10 highest abundances of *narG* gene taxonomies from shotgun metagenome sequencing as detected by GraftM.



Appendix C.3. Heatmap of the 10 highest log10 transformed abundances of *hydB* gene taxonomic annotation from shotgun metagenome sequencing as detected by GraftM.



Appendix C.4. Heatmap of the 25 highest log10 transformed abundances of *acsB* gene taxonomic annotation from shotgun metagenome sequencing as detected by GraftM.



Appendix C.5. Heatmap of the 25 highest log10 transformed abundances of *aprA* gene taxonomies from shotgun metagenome sequencing as detected by GraftM.

Appendix C.6. Heatmap of the 25 highest log10 transformed abundances of *mcrA* gene taxonomies from shotgun metagenome sequencing as detected by GraftM.

Appendix D: Ravenglass and Irish Sea Mud Patch: AOV and Tukey HSD Tables

Appendix D.1. Ravenglass depth core: Summary of *p* values derived from one-way analysis of variance (AOV) and *post hoc* Tukey HSD of values across all and between pairs of sample groups, for (A). Volumes of DNA extracted (ng/μL), and (B). Volume of TOC as percentage of weight. Dash denotes *p* values between 0.05 – 0.1, * denotes 0.01 – 0.05, ** denotes 0.001 – 0.01, and *** denotes < 0.001.

A. DNA (ng/μL) AOV p-value = < 0.001 ***				B. TOC (w/t %) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.003 **			**2**	< 0.001 ***		
3	< 0.001 ***	0.013 *		**3**	< 0.001 ***	0.006 **	
4	< 0.001 ***	0.019 *	1	**4**	< 0.001 ***	0.004 **	0.981

Appendix D.2. Mud Patch Depth Core: Summary of *p* values derived from one-way analysis of variance (AOV) and *post hoc* Tukey HSD of values across all and between pairs of sample groups, for (A). Concentration of DNA extracted (ng/μL), and (B). Volume of TOC as percentage of weight.

A. DNA (ng/μL) AOV p-value = 0.230					B. TOC (w/t %) AOV p-value = 0.063 -				
Group	**1**	**2**	**3**	**4**	**Group**	**1**	**2**	**3**	**4**
2	1				**2**	1			
3	0.999	0.992			**3**	0.071 -	0.071 -		
4	0.451	0.362	0.603		**4**	0.786	0.786	0.454	
5	0.532	0.477	0.696	1	**5**	0.857	0.857	0.591	1

Appendix D.3. Ravenglass Depth Core: Summary of *p* values from AOV and Tukey HSD on 'end point' PCR functional gene band intensities as calculated in BioNumerics, across and between sample groups.

A. Archaeal ammonia monooxygenase AOV p-value = 0.423				B. Bacterial ammonia monooxygenase AOV p-value = 0.001 **			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.786			**2**	0.031		
3	0.532	0.971		**3**	0.831	0.145	
4	0.999	0.753	0.512	**4**	0.164	< 0.001 ***	0.039 *
C. Nitrite reductase – *nirK* AOV p-value = < 0.001 *				**D. Nitrite reductase – *nirS* AOV p-value = 0.103**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.992			**2**	0.203		
3	< 0.001 ***	< 0.001 ***		**3**	0.967	0.396	
4	< 0.001 ***	< 0.001 ***	0.934	**4**	0.945	0.098 -	0.761

E. Dissimilatory sulfite reductase – *dsrA* AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	0.016 *		
3	< 0.001 ***	0.023 *	
4	< 0.001 ***	0.006 **	0.813

Appendix D.4. Mud Patch Depth Core: Summary of *p* values from AOV and Tukey HSD on 'end point' PCR functional gene band intensities as calculated in BioNumerics, across and between sample groups.

A. Archaeal ammonia monooxygenase
AOV p-value = < 0.001 *

Group	1	2	3	4
2	0.232			
3	0.383	0.997		
4	0.193	0.002 **	0.005 **	
5	0.266	0.007 **	0.013 *	1

B. Bacterial ammonia monooxygenase
AOV p-value = 0.449

Group	1	2	3	4
2	1			
3	1	1		
4	0.570	0.513	0.503	
5	0.978	0.963	0.961	0.945

C. Nitrite reductase – *nirK*
AOV p-value = 0.152

Group	1	2	3	4
2	0.999			
3	0.993	1		
4	0.231	0.337	0.423	
5	0.443	0.569	0.656	1

D. Nitrite reductase – *nirS*
AOV p-value = 0.002 **

Group	1	2	3	4
2	0.010 *			
3	0.937	0.051 -		
4	0.812	0.001 **	0.377	
5	1	0.024 *	0.933	0.920

E. Dissimilatory sulfite reductase – *dsrA*
AOV p-value = 0.044 *

Group	1	2	3	4
2	0.628			
3	0.381	0.992		
4	0.466	0.999	1	
5	0.021 *	0.213	0.374	0.308

Appendix D.5. Summary of *p* values derived from analysis of variance (AOV) of three alpha diversity indices between all respective sample groups within the Ravenglass and Mud Patch depth cores, and from Tukey HSD of alpha diversity indices between each sample group within a core.

Ravenglass

A. Shannon Species Richness
AOV p-value = 0.009 **

Group	1	2	3
2	0.446		
3	0.893	0.159	
4	0.083	0.005 **	0.254

C. Faith Phylogenetic Diversity
AOV p-value = 0.063 -

Group	1	2	3
2	0.778		
3	0.990	0.606	
4	0.208	0.043 *	0.317

E. Pielou Species Evenness
AOV p-value = 0.028 *

Group	1	2	3
2	0.052 -		
3	0.583	0.425	
4	0.986	0.038 *	0.431

Irish Sea Mud Patch

B. Shannon Species Richness
AOV p-value = 0.177

Group	1	2	3	4
2	0.168			
3	0.579	0.941		
4	0.392	0.979	0.999	
5	0.280	1	0.962	0.988

D. Faith Phylogenetic Diversity
AOV p-value = 0.587

Group	1	2	3	4
2	0.763			
3	0.636	0.998		
4	0.783	1	0.997	
5	0.648	0.996	1	0.995

F. Pielou Species Evenness
AOV p-value = 0.066 -

Group	1	2	3	4
2	0.071 -			
3	0.508	0.815		
4	0.133	0.997	0.940	
5	0.164	1	0.903	1

Appendix D.6. Ravenglass depth core: Summary of *p* values derived from PERMANOVA across sample groups and pairwise PERMANOVA between pairs of sample groups of sample distances. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac; (C). Unweighted UniFrac distances.

A. Bray Curtis Similarity PERMANOVA p-value = 0.001 **			
Group	**1**	**2**	**3**
2	0.007 **		
3	0.008 **	0.020 *	
4	0.010 *	0.008 **	0.152

B. Weighted UniFrac PERMANOVA p-value = 0.001 **			
Group	**1**	**2**	**3**
2	0.013 *		
3	0.008 **	0.008 **	
4	0.003 **	0.010 *	0.279

C. Unweighted UniFrac PERMANOVA p-value = 0.001			
Group	**1**	**2**	**3**
2	0.007 **		
3	0.009 **	0.032 *	
4	0.009 **	0.021 *	0.036 *

Appendix D.7. Mud patch depth core: Summary of *p* values derived from PERMANOVA across sample groups and pairwise PERMANOVA between pairs of sample groups of sample distances. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac; (C). Unweighted UniFrac distances.

A. Bray Curtis Similarity AOV p-value = 0.002 **				
Group	**1**	**2**	**3**	**4**
2	0.008 **			
3	0.140	0.059 -		
4	0.003 **	0.008 **	0.513	
5	0.025 *	0.023 *	0.215	0.837

B. Weighted UniFrac AOV p-value = 0.084 -				
Group	**1**	**2**	**3**	**4**
2	0.095 -			
3	0.249	0.191		
4	0.101	0.008 **	0.549	
5	0.107	0.016 *	0.736	0.768

C. Unweighted UniFrac AOV p-value = 0.001 **				
Group	**1**	**2**	**3**	**4**
2	0.005 **			
3	0.011 *	0.046 *		
4	0.011 *	0.018 *	0.854	
5	0.020 *	0.014 *	0.656	0.787

Appendix D.8. Ravenglass Depth Core: Summary of *p* values derived from AOV and *post hoc* Tukey HSD of ASV abundance collapsed at genus level across and between the four depth sample groups.

A. *Woeseia* spp. AOV p-value = < 0.001 ***				**B. *Ardenticatenales* AOV p-value = 0.881**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.650			**2**	0.983		
3	0.002 **	0.018 *		**3**	0.999	0.996	
4	< 0.001 ***	0.003 **	0.648	**4**	0.866	0.973	0.922
C. δ-proteobacteria MBNT-15 AOV p-value = < 0.001 ***				**D. *Anaerolineaceae* AOV p-value = 0.449**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.935			**2**	0.743		
3	0.002 **	0.007 **		**3**	0.960	0.459	
4	0.001 **	0.002 **	0.818	**4**	0.978	0.545	1
E. *Gemmatimonadetes* PAUC43f MBG AOV p-value = < 0.001 ***				**F. *Pirellulaceae* AOV p-value = 0.001 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.206			**2**	0.693		
3	0.001 **	0.036 *		**3**	0.007 **	0.061 -	
4	< 0.001 ***	0.002 **	0.319	**4**	0.002 **	0.016 *	0.824
G. γ-*proteobacteria* AOV p-value = 0.001 **				**H. *Pirellulaceae* Pir4 AOV p-value = 0.002 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.618			**2**	0.320		
3	0.006 **	0.064 -		**3**	0.003 **	0.082 -	
4	0.003 **	0.026 *	0.921	**4**	0.004 **	0.099 -	1
I. *Thermoanaerobaculaceae* SG10 AOV p-value = < 0.001 ***				**J. *Ardenticatenales* (2) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.446			**2**	0.369		
3	0.008 **	0.140		**3**	0.001 **	0.035 *	
4	0.001 **	0.019 *	0.623	**4**	< 0.001 ***	0.009 **	0.815
K. *Desulfobulbaceae* (1) AOV p-value = < 0.001 ***				**L. *Chloroflexi* KD4-96 AOV p-value = 0.054 -**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.191		
3	< 0.001 ***	0.005 **		**3**	0.293	0.992	
4	< 0.001 ***	0.002 **	0.855	**4**	0.041 *	0.759	0.607
M. *Sandaracinaceae* AOV p-value = 0.236				**N. *Dadabacteriales* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.999			**2**	0.290		
3	0.474	0.394		**3**	< 0.001 ***	0.004 **	
4	0.492	0.416	1.000	**4**	< 0.001 ***	0.005 **	0.997
O. *Desulfobulbaceae* (2) AOV p-value = < 0.001 ***				**P. *Saprospiraceae* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.001 **			**2**	0.012 *		
3	< 0.001 ***	0.304		**3**	< 0.001 ***	0.047 *	
4	< 0.001 ***	0.234	0.992	**4**	< 0.001 ***	0.025 *	0.957
Q. *Prolixibacteraceae* AOV p-value = 0.001 **				**R. *Methyloceanibacter* spp. AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.015 *			**2**	0.437		
3	0.002 **	0.777		**3**	< 0.001 ***	0.003 **	
4	0.004 **	0.806	1	**4**	< 0.001 ***	0.003 **	0.979
S. *Nitrosopumilus* spp. AOV p-value = 0.022 *				**T. *Bathyarchaeia* AOV p-value = 0.163**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.054 -			**2**	0.273		
3	0.386	0.638		**3**	0.156	0.984	
4	0.025 *	0.938	0.356	**4**	0.468	0.991	0.921

Appendix D.9. Mud Patch Depth Core: Summary of *p* values derived from AOV and *post hoc* Tukey HSD of ASV abundance collapsed at genus level across and between the four depth sample groups.

A. *Planococcaceae*
AOV p-value = 0.599

Group	1	2	3	4
2	0.674			
3	0.999	0.838		
4	0.835	0.998	0.941	
5	0.767	1	0.885	0.999

B. *Desulfobulbaceae* (1)
AOV p-value = 0.002 **

Group	1	2	3	4
2	0.002 **			
3	0.774	0.037 *		
4	0.981	0.006 **	0.965	
5	0.996	0.016 *	0.962	1

C. *Flavobacteriaceae*
AOV p-value = 0.025 *

Group	1	2	3	4
2	0.120			
3	0.066 -	0.989		
4	0.029 *	0.941	0.999	
5	0.086 -	0.983	1	1

D. γ-*proteobacteria* B2M28
AOV p-value = 0.005 **

Group	1	2	3	4
2	0.003 **			
3	0.049 *	0.773		
4	0.051 -	0.609	1	
5	0.096 -	0.765	1	1

E. *Psychrobacter* spp.
AOV p-value = 0.234

Group	1	2	3	4
2	0.261			
3	0.325	1		
4	0.402	0.998	0.999	
5	0.906	0.867	0.894	0.954

F. γ-*proteobacteria* (1)
AOV p-value = 0.012 *

Group	1	2	3	4
2	0.012 *			
3	0.973	0.060 -		
4	0.970	0.041 *	1	
5	0.998	0.059 -	0.999	0.999

G. *Syntrophobacteraceae*
AOV p-value = 0.010 *

Group	1	2	3	4
2	0.378			
3	0.420	1		
4	0.008 **	0.274	0.338	
5	0.036 *	0.513	0.568	1

H. *Anaerolineaceae*
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	0.095 -			
3	0.007 **	0.564		
4	< 0.001 ***	0.145	0.925	
5	< 0.001 ***	0.041 *	0.481	0.854

I. γ-*proteobacteria* (2)
AOV p-value = 0.228

Group	1	2	3	4
2	0.214			
3	0.952	0.636		
4	0.962	0.530	1	
5	1	0.324	0.965	0.974

J. *Desulfobulbaceae* (2)
AOV p-value = 0.001 **

Group	1	2	3	4
2	0.003 **			
3	0.014 *	0.980		
4	0.001 **	0.993	0.875	
5	0.025 *	0.987	1	0.910

K. *Lutimonas* spp.
AOV p-value = 0.023 *

Group	1	2	3	4
2	0.064 -			
3	0.063 -	1		
4	0.039 *	0.999	1	
5	0.068 -	0.996	1	1

L. *Woeseia* spp.
AOV p-value = 0.792

Group	1	2	3	4
2	0.984			
3	0.822	0.977		
4	0.995	1	0.952	
5	0.841	0.977	1	0.954

M. *Thermoanaerobaculaceae* subgroup 10
AOV p-value = 0.017 *

Group	1	2	3	4
2	0.051 -			
3	0.178	0.984		
4	0.015 *	0.969	0.803	
5	0.084 -	1	0.975	0.995

N. *Methyloceanibacter* spp.
AOV p-value = 0.003 **

Group	1	2	3	4
2	0.244			
3	0.300	1		
4	0.065 -	0.940	0.948	
5	0.001 **	0.050 *	0.063 -	0.166

O. δ-*proteobacteria* SEEP-SRB1
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	0.177			
3	0.036 *	0.852		
4	< 0.001 ***	0.041 *	0.321	
5	< 0.001 ***	0.029 *	0.195	0.975

P. *Bathyarchaeia*
AOV p-value = 0.004 **

Group	1	2	3	4
2	1			
3	0.521	0.638		
4	0.193	0.269	0.977	
5	0.004 **	0.005 **	0.090 -	0.182

Appendix D.10. Ravenglass Depth Core: Summary of p values from AOV and post hoc Tukey HSD on Tax4Fun inferred gene abundances of the genes targeted by 'end point' PCR across and between the four sample resolution groups.

A. Ammonia monooxygenase – *amoA*
AOV p-value = 0.015 *

Group	1	2	3
2	0.039 *		
3	0.227	0.747	
4	0.016 *	0.922	0.419

B. Nitrite reductase – *nirK*
AOV p-value = 0.007 **

Group	1	2	3
2	0.014 *		
3	0.122	0.668	
4	0.011 *	0.987	0.511

C. Nitrite reductase – *nirS*
AOV p-value = 0.006 **

Group	1	2	3
2	0.305		
3	0.027 *	0.526	
4	0.005 **	0.136	0.740

D. Dissimilatory sulfite reductase – *dsrA*
AOV p-value = < 0.001 ***

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.044 *	
4	< 0.001 ***	0.006 **	0.660

E. Methyl Coenzyme M Reductase – *mcrA*
AOV p-value = 0.060 -

Group	1	2	3
2	0.918		
3	0.382	0.744	
4	0.053 -	0.155	0.581

Appendix D.11. Ravenglass Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of nitrogen metabolism genes across and between the four sample resolution groups.

A. Nitrate reductase – *narG*
AOV p-value = 0.005 **

Group	1	2	3
2	0.691		
3	0.256	0.037 *	
4	0.051 -	0.007 **	0.483

B. Nitrate reductase – *napA*
AOV p-value = < 0.001 ***

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.052 -	
4	< 0.001 ***	0.002 **	0.284

C. Nitric oxide reductase – *norB*
AOV p-value = 0.002 **

Group	1	2	3
2	0.056 -		
3	0.993	0.092 -	
4	0.184	0.001 **	0.119

D. Nitrous-oxide reductase – *nosZ*
AOV p-value = 0.006 **

Group	1	2	3
2	1		
3	0.107	0.103	
4	0.015 *	0.014 *	0.642

E. Nitrite reductase – *nrfA*
AOV p-value = 0.012 *

Group	1	2	3
2	0.012 *		
3	0.069 -	0.804	
4	0.047 *	0.960	0.984

F. Nitrogenase – *nifD*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.010 *		
3	0.002 **	0.835	
4	< 0.001 ***	0.034 *	0.142

Appendix D.12. Ravenglass Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of genes across and between the four sample resolution groups.

A. Acetyl-CoA synthase – *acsB*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.147	
4	< 0.001 ***	0.068 -	0.940

B. Adenylylsulfate reductase – *aprA*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.042 *	
4	< 0.001 ***	0.006 **	0.636

C. Methane monooxygenase – *mmoX*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.902	
4	0.003 **	0.085 -	0.251

D. Methanol dehydrogenase – *mxaF*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.565	
4	0.002 **	0.979	0.388

E. NiFe hydrogenase – *hydB*
AOV p-value = 0.012 *

Group	1	2	3
2	0.016 *		
3	0.785	0.095 -	
4	1	0.024 *	0.814

F. Methylene-tetrahydromethanopterin dehydrogenase – *mtdB*
AOV p-value = < 0.001 *

Group	1	2	3
2	0.003 **		
3	< 0.001 ***	0.015 *	
4	< 0.001 ***	0.063 -	0.944

G. Heterodisulfide reductase – *hdrA*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.305	
4	< 0.001 ***	0.096 -	0.845

H. Anaerobic carbon-monoxide dehydrogenase – *cooS*
AOV p-value = < 0.001 *

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.199	
4	< 0.001 ***	0.061 -	0.853

Appendix D.13. Mud Patch Depth Core: Summary of p values from AOV and post hoc Tukey HSD on Tax4Fun inferred gene abundances of the genes targeted by 'end point' PCR across and between the four sample resolution groups.

A. Ammonia monooxygenase – *amoA*
AOV p-value = 0.457

Group	1	2	3	4
2	0.720			
3	0.494	0.990		
4	0.998	0.866	0.655	
5	0.832	1	0.991	0.929

B. Nitrite reductase – *nirK*
AOV p-value = 0.136

Group	1	2	3	4
2	0.340			
3	0.570	0.998		
4	0.127	0.971	0.898	
5	0.230	0.984	0.937	1

C. Nitrite reductase – *nirS*
AOV p-value = 0.034 *

Group	1	2	3	4
2	0.754			
3	0.294	0.891		
4	0.058 -	0.419	0.936	
5	0.057 -	0.335	0.823	0.994

D. Dissimilatory sulfite reductase – *dsrA*
AOV p-value = < 0.001 *

Group	1	2	3	4
2	< 0.001 ***			
3	< 0.001 ***	0.365		
4	< 0.001 ***	0.036 *	0.778	
5	< 0.001 ***	0.042 *	0.670	0.997

E. Methyl Coenzyme M reductase – *mcrA*
AOV p-value = 0.347

Group	1	2	3	4
2	1			
3	0.410	0.393		
4	1	1	0.408	
5	1	1	0.520	1

Appendix D.14. Mud Patch Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of nitrogen metabolism genes across and between the four sample resolution groups.

A. Nitrate reductase – *narG*
AOV p-value = 0.001 **

Group	1	2	3	4
2	0.205			
3	0.014 *	0.548		
4	0.002 **	0.162	0.950	
5	0.007 **	0.273	0.968	1

B. Nitrate reductase – *napA*
AOV p-value = 0.060 -

Group	1	2	3	4
2	0.217			
3	0.117	0.987		
4	0.059 -	0.946	1	
5	0.231	0.999	1	0.996

C. Nitric oxide reductase – *norB*
AOV p-value = 0.080 -

Group	1	2	3	4
2	1			
3	0.760	0.768		
4	0.146	0.151	0.785	
5	0.335	0.342	0.923	1

D. Nitrous-oxide reductase – *nosZ*
AOV p-value = 0.199

Group	1	2	3	4
2	0.326			
3	0.222	0.996		
4	0.288	1	0.998	
5	0.507	1	0.994	1

E. Nitrite reductase – *nrfA*
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	< 0.001 ***			
3	< 0.001 ***	0.991		
4	< 0.001 ***	0.985	1	
5	< 0.001 ***	0.999	1	1

F. Nitrogenase – *nifD*
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	< 0.001 ***			
3	< 0.001 ***	0.332		
4	< 0.001 ***	0.068 -	0.936	
5	< 0.001 ***	0.190	0.987	1

Appendix D.15. Mud Patch Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of genes across and between the four sample resolution groups.

A. Acetyl-CoA synthase – *acsB*
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	0.005 **			
3	< 0.001 ***	0.228		
4	< 0.001 ***	0.008 **	0.543	
5	< 0.001 ***	0.009 **	0.403	0.990

B. Adenylylsulfate reductase – *aprA*
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	< 0.001 ***			
3	< 0.001 ***	0.243		
4	< 0.001 ***	0.030 *	0.869	
5	< 0.001 ***	0.036 *	0.768	0.997

C. Methane monooxygenase – *mmoX*
AOV p-value = 0.632

Group	1	2	3	4
2	0.737			
3	0.999	0.891		
4	1	0.792	1	
5	0.815	1	0.924	0.857

D. Methanol dehydrogenase – *mxaF*
AOV p-value = 0.735

Group	1	2	3	4
2	1			
3	0.970	0.943		
4	0.862	0.800	0.998	
5	1	1	0.945	0.831

E. NiFe hydrogenase – *hydB*
AOV p-value = 0.023 *

Group	1	2	3	4
2	0.420			
3	0.287	0.995		
4	0.023 *	0.482	0.757	
5	0.049 *	0.567	0.795	1

F. Methylene-tetrahydromethanopterin dehydrogenase – *mtdB*
AOV p-value = 0.858

Group	1	2	3	4
2	1			
3	0.997	1		
4	0.982	0.994	1	
5	0.972	0.946	0.897	0.816

G. Heterodisulfide reductase – *hdrA*
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	0.033 *			
3	0.003 **	0.638		
4	< 0.001 ***	0.068 -	0.682	
5	< 0.001 ***	0.092 -	0.650	0.999

H. Anaerobic carbon-monoxide dehydrogenase – *cooS*
AOV p-value = < 0.001 ***

Group	1	2	3	4
2	< 0.001 ***			
3	< 0.001 ***	0.286		
4	< 0.001 ***	0.007 **	0.435	
5	< 0.001 ***	0.007 **	0.291	0.983

Appendix D.16. Ravenglass depth core: Summary of *p* values derived from one-way analysis of variance (AOV) and *post hoc* Tukey HSD of a range of geochemical concentrations and radionuclide activities across and between pairs of sample groups. (A). ICP-AES porewater manganese (Mn) concentration in μM; (B). ICP porewater iron (Fe) concentrations in μM; (C). Ion chromatograph sulfate (SO_4^{2-}) concentrations in mM; (D). XRF Mn normalised by Al volumes (%); (E). XRF Fe normalised by Al volumes (%); (F). XRF ^{241}Am total activity (Bq/Kg); (G). XRF ^{137}Cs total activity (Bq/Kg); (H). α-spectrometer $^{239/240}$Pu total activity (Bq/Kg); (I). Scintillation spectrometer ^{238}Pu total activity (Bq/Kg).

A. Mn (μM) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	< 0.001 ***		
3	< 0.001 ***	1	
4	< 0.001 ***	0.998	1

B. Fe (μM) AOV p-value = 0.023 *			
Group	**1**	**2**	**3**
2	0.508		
3	0.083 -	0.643	
4	0.022 *	0.242	0.830

C. SO_4^{2-} (mM) AOV p-value = 0.580			
Group	**1**	**2**	**3**
2	0.568		
3	0.943	0.872	
4	0.731	0.997	0.955

D. Mn/Al (%) AOV p-value = 0.020 *			
Group	**1**	**2**	**3**
2	0.852		
3	0.016 *	0.072	
4	0.421	0.845	0.350

E. Fe/Al (%) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	0.799		
3	0.979	0.953	
4	0.003 **	< 0.001 ***	0.002 **

F. ^{241}Am (Bq/Kg) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	0.839		
3	0.007 **	0.036	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

G. ^{137}Cs (Bq/Kg) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	0.906		
3	0.009 **	0.035	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

H. $^{239/240}$Pu (Bq/Kg) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	0.940		
3	0.005 **	0.015	
4	< 0.001 ***	< 0.001 ***	0.030 *

I. ^{238}Pu (Bq/Kg) AOV p-value = < 0.001 ***			
Group	**1**	**2**	**3**
2	0.976		
3	0.21 *	0.046 *	
4	< 0.001 ***	< 0.001 ***	0.167

Appendix D.17. Mud patch depth core: Summary of *p* values derived from one-way analysis of variance (AOV) and *post hoc* Tukey HSD of a range of geochemical concentrations and radionuclide activities across and between pairs of sample groups. (A). ICP-AES porewater manganese (Mn) concentration in μM; (B). ICP porewater iron (Fe) concentrations in μM; (C). Ion chromatograph sulfate (SO_4^{2-}) concentrations in mM; (D). XRF Mn normalised by Al volumes (%); (E). XRF Fe normalised by Al volumes (%); (F). XRF ^{241}Am total activity (Bq/Kg); (G). XRF ^{137}Cs total activity (Bq/Kg); (H). α-spectrometer $^{239/240}$Pu total activity (Bq/Kg); (I). Scintillation spectrometer ^{238}Pu total activity (Bq/Kg).

A. Mn (μM) AOV p-value = < 0.001 ***					**B. Fe (μM) AOV p-value = 0.266**				
Group	**1**	**2**	**3**	**4**	**Group**	**1**	**2**	**3**	**4**
2	< 0.001 ***				**2**	0.452			
3	< 0.001 ***	0.198			**3**	0.684	0.995		
4	< 0.001 ***	0.054 -	0.952		**4**	0.982	0.768	0.934	
5	< 0.001 ***	0.039 *	0.765	0.979	**5**	0.989	0.326	0.507	0.876
C. SO_4^{2-} (μM) AOV p-value = < 0.001 ***					**D. Mn/Al (%) AOV p-value = < 0.001 *****				
Group	**1**	**2**	**3**	**4**	**Group**	**1**	**2**	**3**	**4**
2	< 0.001 ***				**2**	0.003 **			
3	0.196	< 0.001 ****			**3**	0.002 **	0.999		
4	0.118	< 0.001 ***	0.998		**4**	< 0.001 ***	0.915	0.977	
5	0.726	< 0.001 ***	0.946	0.859	**5**	< 0.001 ***	0.515	0.650	0.903
E. Fe/Al (%) AOV p-value = 0.026 *					**F. ^{241}Am (Bq/Kg) AOV p-value = 0.002 ****				
Group	**1**	**2**	**3**	**4**	**Group**	**1**	**2**	**3**	**4**
2	0.614				**2**	0.676			
3	0.171	0.883			**3**	0.049 *	0.456		
4	0.030 *	0.380	0.885		**4**	0.021 *	0.255	0.993	
5	0.067 -	0.498	0.918	1	**5**	0.002 **	0.020 *	0.304	0.490
G. ^{137}Cs (Bq/Kg) AOV p-value = < 0.001 ***					**H. $^{239/240}$Pu (Bq/Kg) AOV p-value = 0.005 ****				
Group	**1**	**2**	**3**	**4**	**Group**	**1**	**2**	**3**	**4**
2	0.095 -				**2**	0.406			
3	< 0.001 ***	0.021 *			**3**	0.052 -	0.748		
4	< 0.001 ***	0.050 *	0.993		**4**	0.003 **	0.117	0.653	
5	< 0.001 ***	0.035 *	1	0.977	**5**	0.091 -	0.772	1	0.819

I. ^{238}Pu (Bq/Kg) AOV p-value = 0.004 **				
Group	**1**	**2**	**3**	**4**
2	0.438			
3	0.078 -	0.824		
4	0.002 **	0.071 -	0.413	
5	0.075 -	0.689	0.995	0.777

Appendix E. Loch Etive: AOV and Tukey HSD Tables

Appendix E.1. Summary of *p* values from one-way analysis (AOV) of variance and *post hoc* Tukey HSD of extracted DNA volumes across and between sample groups from: (A). Shallow Basin depth core, and; (B). Deep Basin depth core. Dash denotes *p* values between 0.05 – 0.1, * denotes 0.01 – 0.05, ** denotes 0.001 – 0.01, and *** denotes < 0.001.

A. Shallow Basin AOV p-value = < 0.001 * **				**B. Deep Basin AOV p-value = < 0.001 *** **			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.216			**2**	0.166		
3	< 0.001 ***	< 0.001 ***		**3**	0.035 *	0.697	
4	< 0.001 ***	< 0.001 ***	0.889	**4**	< 0.001 ***	0.007 **	0.083 -

Appendix E.2. Shallow Basin: Summary of *p* values from AOV and Tukey HSD on 'end point' PCR functional gene band intensities as calculated in BioNumerics, across and between sample groups.

A. Archaeal ammonia monooxygenase AOV p-value = 0.078 -				**B. Bacterial ammonia monooxygenase AOV p-value = 0.007 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.800			**2**	0.933		
3	0.968	0.453		**3**	0.054 -	0.005 **	
4	0.390	0.055 -	0.598	**4**	0.616	0.222	0.458
C. Nitrite reductase – *nirK* AOV p-value = < 0.001 * **				**D. Nitrite reductase – *nirS* AOV p-value = 0.014 ***			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.082 -			**2**	0.708		
3	0.329	< 0.001 ***		**3**	0.073 -	0.245	
4	0.588	0.004 **	0.977	**4**	0.022 *	0.069 -	0.861

Appendix E.3. Deep Basin: Summary of *p* values from AOV and Tukey HSD on 'end point' PCR functional gene band intensities as calculated in BioNumerics, across and between sample groups.

A. Archaeal ammonia monooxygenase AOV p-value = 0.454				**B. Bacterial ammonia monooxygenase AOV p-value = < 0.001 *** **			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.968			**2**	< 0.001 ***		
3	0.664	0.825		**3**	< 0.001 ***	0.748	
4	0.990	0.833	0.423	**4**	< 0.001 ***	0.069 -	0.405
C. Nitrite reductase – *nirK* AOV p-value = < 0.001 * **				**D. Nitrite reductase – *nirS* AOV p-value = 0.003 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.010 *		
3	< 0.001 ***	0.996		**3**	0.004 **	0.892	
4	< 0.001 ***	0.989	0.957	**4**	0.005 **	0.887	1

Appendix E.4. Summary of p-values derived from AOV of three alpha diversity indices between all respective sample groups within each core, and Tukey HSD between sample groups of each core.

Shallow Basin				Deep Basin			
A. Shannon Species Richness AOV p-value = 0.258				**B. Shannon Species Richness AOV p-value = 0.060 -**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.691			**2**	0.956		
3	0.992	0.825		**3**	0.840	0.419	
4	0.845	0.209	0.670	**4**	0.288	0.396	0.040 *
C. Faith Phylogenetic Diversity AOV p-value = 0.150				**D. Faith Phylogenetic Diversity AOV p-value = 0.004 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.242			**2**	0.260		
3	0.162	0.961		**3**	0.819	0.685	
4	0.223	0.983	1	**4**	0.004 **	0.072 -	0.013 *
E. Pielou Species Evenness AOV p-value = < 0.001 ***				**F. Pielou Species Evenness AOV p-value = 0.037 ***			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.131			**2**	0.032 *		
3	0.003 **	0.091		**3**	0.379	0.475	
4	< 0.001 ***	0.002 **	0.312	**4**	0.084 -	0.994	0.714

Appendix E.5. Shallow Basin Depth Core: Summary of *p* values derived from PERMANOVA across sample groups and pairwise PERMANOVA between pairs of sample groups of sample distances. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac; (C). Unweighted UniFrac distances.

A. Bray Curtis dissimilarity PERMANOVA p-value = 0.001 **				**B. Weighted UniFrac PERMANOVA p-value = 0.001 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.011 *			**2**	0.006 **		
3	0.010 *	0.010 *		**3**	0.019 *	0.006 **	
4	0.027 *	0.010 *	0.013 *	**4**	0.031 *	0.010 *	0.019 *
C. Unweighted UniFrac PERMANOVA p-value = 0.001							
Group	**1**	**2**	**3**				
2	0.006 **						
3	0.022 *	0.046 *					
4	0.046 *	0.006 **	0.106				

Appendix E.6. Deep Basin Depth Core: Summary of *p* values derived from PERMANOVA across sample groups and pairwise PERMANOVA between pairs of sample groups of sample distances. (A). Bray Curtis dissimilarity; (B). Weighted UniFrac; (C). Unweighted UniFrac distances.

A. Bray Curtis Dissimilarity PERMANOVA p-value = 0.001 **				**B. Weighted UniFrac PERMANOVA p-value = 0.001 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.008 **			**2**	0.010 *		
3	0.012 *	0.012 *		**3**	0.014 *	0.008 **	
4	0.011 *	0.012 *	0.012 *	**4**	0.008 **	0.006 **	0.010 *
C. Unweighted UniFrac PERMANOVA p-value = 0.001							
Group	**1**	**2**	**3**				
2	0.008 **						
3	0.010 *	0.006 **					
4	0.010 *	0.006 **	0.010 *				

Appendix E.7. Shallow Basin: Summary of *p* values derived from AOV and *post hoc* Tukey HSD of ASV abundance collapsed to genus level across and between the four sample resolution groups.

A. Uncultured *Tepidibacter* spp.
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.148		
3	0.005 **	0.143	
4	< 0.001 ***	< 0.001 ***	0.006 **

B. *Clostridiisalibacter* spp.
AOV p-value = 0.044 *

Group	1	2	3
2	0.793		
3	0.292	0.027 *	
4	0.988	0.580	0.461

C. Uncultured *Anaerolineaceae*
AOV p-value = 0.005 **

Group	1	2	3
2	0.013 *		
3	0.004 **	0.734	
4	0.044 *	0.999	0.755

D. Uncultured *Sandaracinaceae*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.154		
3	0.160	0.997	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

E. Uncultured *Thermoanaerobaculaceae* SG23
AOV p-value = 0.004 **

Group	1	2	3
2	0.005 **		
3	0.175	0.332	
4	0.822	0.037 *	0.607

F. Uncultured *Desulfobulbaceae*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.011 *		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.618

G. *Alkaliphilus* spp.
AOV p-value = 0.075 -

Group	1	2	3
2	0.218		
3	0.053 -	0.672	
4	0.262	0.998	0.851

H. Uncultured *Bacteriodetes* BD2-2
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.057 -		
3	0.025 *	0.863	
4	0.175	< 0.001 ***	< 0.001 ***

I. Uncultured *Gammaproteobacteria* B2M28
AOV p-value = 0.006 **

Group	1	2	3
2	0.707		
3	0.844	0.997	
4	0.087	0.004 **	0.013 *

J. *Fusibacter* spp.
AOV p-value = 0.059 -

Group	1	2	3
2	0.672		
3	0.109	0.386	
4	0.097 -	0.333	0.997

K. *Filomicrobium* spp.
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.858		
3	0.001 **	0.001 **	
4	< 0.001 ***	< 0.001 ***	0.065 -

L. Uncultured δ-*proteobacteria* NB1-j
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.007 **		
3	0.328	0.208	
4	0.489	< 0.001 ***	0.021 *

M. Uncultured *Bathyarchaeia*
AOV p-value = < 0.001 ***

Group	1	2	3
2	< 0.001 ***		
3	0.004 **	0.384	
4	0.007 **	0.435	1

N. Uncultured *Desulfobacteraceae* Sva0081
AOV p-value = 0.001 **

Group	1	2	3
2	0.082		
3	0.996	0.034 *	
4	0.288	0.001 **	0.342

O. *Psychrilyobacter* spp.
AOV p-value = 0.023 *

Group	1	2	3
2	0.725		
3	0.051 -	0.162	
4	0.993	0.537	0.029 *

P. *Thiogranum* spp.
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.335		
3	< 0.001 ***	0.005 **	
4	< 0.001 ***	< 0.001 ***	0.108

Appendix E.8. Deep Basin Depth Core: Summary of *p* values derived from AOV and *post hoc* Tukey HSD of ASV abundance collapsed to genus level across and between the four sample resolution groups.

A. Uncultured *Anaerolineaceae*
AOV p-value = < 0.001 ***

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.826	
4	< 0.001 ***	0.492	0.182

B. Uncultured *Sandaracinaceae*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.003 **		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

C. Uncultured *Bacteriodetes BD2-2*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.012 *		
3	0.699	< 0.001 ***	
4	0.004 **	< 0.001 ***	0.018 *

D. Uncultured *Thermoanaerobaculaceae* SG23
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.177		
3	0.507	0.003 **	
4	0.026 *	< 0.001 ***	0.241

E. Uncultured *Clostridiaceae*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.019 *		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.947

F. Uncultured *Bathyarchaeia*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.307		
3	0.003 **	0.041 *	
4	< 0.001 ***	< 0.001 ***	0.001 **

G. Uncultured *Syntrophobacteraceae*
AOV p-value = 0.086 -

Group	1	2	3
2	0.185		
3	0.099 -	0.946	
4	0.827	0.598	0.364

H. Uncultured *Gammaproteobacteria*
AOV p-value = < 0.001 ***

Group	1	2	3
2	< 0.001 ***		
3	< 0.001 ***	0.030 *	
4	< 0.001 ***	0.017 *	0.977

I. *Ignavibacteriales* PHOS-HE36
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.007 **		
3	< 0.001 ***	0.065 -	
4	0.079 -	0.770	0.018 *

J. *Tepidibacter* spp.
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.034 *		
3	< 0.001 ***	0.006 **	
4	0.302	0.690	0.001 **

K. Uncultured *Desulfobulbaceae*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.172		
3	0.001 **	0.022 *	
4	< 0.001 ***	0.002 **	0.567

L. Uncultured *Deltaproteobacteria* NB1-j
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.085 -		
3	0.001 **	0.070 -	
4	0.948	0.190	0.002 **

M. *Desulfatiglans* spp.
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.041 *		
3	< 0.001 ***	0.005 **	
4	< 0.001 ***	< 0.001 ***	0.687

N. *Ignavibacterium* spp.
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.341		
3	0.005 **	0.059 -	
4	0.002 **	0.024 *	0.945

O. Uncultured *Desulfobacteraceae* Sva0081
AOV p-value = 0.003 **

Group	1	2	3
2	0.548		
3	0.994	0.291	
4	0.093 -	0.002 **	0.093 -

P. Uncultured *Desulfobacteraceae* SEEP-SRB1
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.003 **		
3	< 0.001 ***	0.070 -	
4	0.040 *	0.777	0.020 *

Appendix E.9. Shallow Basin Depth Core: Summary of p values from AOV and post hoc Tukey HSD on Tax4Fun inferred gene abundances of the genes targeted by 'end point' PCR across and between the four sample resolution groups.

A. Ammonia monooxygenase – *amoA*
AOV p-value = 0.003 **

Group	1	2	3
2	0.005 **		
3	0.039 *	0.853	
4	0.858	0.031 *	0.183

B. Nitrite reductase – *nirK*
AOV p-value = < 0.001 ***

Group	1	2	3
2	< 0.001 ***		
3	0.047 *	0.267	
4	0.954	0.002 **	0.131

C. Nitrite reductase – *nirS*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.025 *		
3	< 0.001 ***	0.021 *	
4	0.006 **	0.595	0.388

D. Dissimilatory sulfite reductase – *dsrA*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.525		
3	0.031 *	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.214

E. Methyl Conenzyme M Reductase – *mcrA*
AOV p-value = 0.280

Group	1	2	3
2	0.998		
3	0.748	0.760	
4	0.360	0.324	0.866

Appendix E.10. Shallow Basin Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of nitrogen metabolism genes across and between the four sample resolution groups.

A. Nitrate reductase – *narG*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.818		
3	0.021 *	< 0.001 ***	
4	0.006 **	< 0.001 ***	0.842

B. Nitrate reductase – *napA*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.999		
3	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.507

C. Nitric oxide reductase – *norB*
AOV p-value = 0.001 **

Group	1	2	3
2	0.911		
3	0.020 *	0.025 *	
4	0.007 **	0.007 **	0.878

D. Nitrous-oxide reductase – *nosZ*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.496		
3	0.008 **	0.051 -	
4	< 0.001 ***	0.002 **	0.399

E. Nitrite reductase – *nrfA*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.634		
3	< 0.001 ***	0.002 **	
4	< 0.001 ***	< 0.001 ***	0.011 *

F. Nitrogenase – *nifD*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.303		
3	0.026 *	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.125

Appendix E.11. Shallow Basin Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of genes across and between the four sample resolution groups.

A. Acetyl-CoA synthase – *acsB*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.139		
3	0.288	0.001 **	
4	0.010 *	< 0.001 ***	0.229

B. Adenylylsulfate reductase – *aprA*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.421		
3	0.054 -	< 0.001 ***	
4	< 0.001 ***	< 0.001 ***	0.146

C. Methane monooxygenase – *mmoX*
AOV p-value = < 0.001 ***

Group	1	2	3
2	1		
3	0.008 **	0.002 **	
4	< 0.001 ***	< 0.001 ***	0.463

D. Methanol dehydrogenase – *mxaF*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.966		
3	0.004 **	0.002 **	
4	< 0.001 ***	< 0.001 ***	0.501

E. NiFe hydrogenase – *hydB*
AOV p-value = 0.931

Group	1	2	3
2	0.912		
3	0.983	0.993	
4	0.989	0.991	1

F. Methylene-tetrahydromethanopterin dehydrogenase – *mtdB*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.573		
3	0.134	0.004 **	
4	0.013 *	< 0.001 ***	0.525

G. Heterodisulfide reductase – *hdrA*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.317		
3	0.165	0.002 **	
4	0.007 **	< 0.001 ***	0.291

H. Anaerobic carbon-monoxide dehydrogenase – *cooS*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.282		
3	0.131	< 0.001 ***	
4	0.005 **	< 0.001 ***	0.288

Appendix E.12. Deep Basin Depth Core: Summary of p values from AOV and post hoc Tukey HSD on Tax4Fun inferred gene abundances of the genes targeted by 'end point' PCR across and between the four sample resolution groups.

A. Ammonia monooxygenase – *amoA*
AOV p-value = 0.505

Group	1	2	3
2	1		
3	0.696	0.606	
4	0.998	0.993	0.533

B. Nitrite reductase – *nirK*
AOV p-value = 0.097 -

Group	1	2	3
2	0.363		
3	0.100	0.729	
4	0.919	0.721	0.254

C. Nitrite reductase – *nirS*
AOV p-value = < 0.001 ***

Group	1	2	3
2	0.213		
3	0.144	0.978	
4	< 0.001 ***	< 0.001 ***	< 0.001 ***

D. Dissimilatory sulfite reductase – *dsrA*
AOV p-value = 0.008 **

Group	1	2	3
2	0.034 *		
3	0.028 *	0.990	
4	0.938	0.085 -	0.067 -

E. Methyl Conenzyme M Reductase – *mcrA*
AOV p-value = 0.336

Group	1	2	3
2	0.497		
3	1	0.431	
4	0.999	0.530	1

Appendix E.13. Deep Basin Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of nitrogen metabolism genes across and between the four sample resolution groups.

A. Nitrate reductase – *narG* AOV p-value = 0.004 **				**B. Nitrate reductase – *napA* AOV p-value = 0.005 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.014 *			**2**	0.132		
3	0.003 **	0.686		**3**	0.495	0.805	
4	0.034 *	0.997	0.645	**4**	0.595	0.004 **	0.093 *
C. Nitric oxide reductase – *norB* AOV p-value = 0.082 -				**D. Nitrous-oxide reductase – *nosZ* AOV p-value = 0.037 ***			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.202			**2**	0.975		
3	0.092 -	0.913		**3**	0.975	1	
4	0.843	0.607	0.325	**4**	0.187	0.040 *	0.054 -
E. Nitrite reductase – *nrfA* AOV p-value = < 0.001 ***				**F. Nitrogenase – *nifD* AOV p-value = 0.002 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.552			**2**	0.007 **		
3	0.163	0.0696		**3**	0.001 **	0.694	
4	< 0.001 ***	< 0.001 ***	0.012 *	**4**	0.206	0.385	0.089 -

Appendix E.14. Deep Basin Depth Core: Summary of *p* values from AOV and *post hoc* Tukey HSD on Tax4Fun inferred gene abundances of genes across and between the four sample resolution groups.

A. Acetyl-CoA synthase – *acsB* AOV p-value = 0.004 **				**B. Adenylylsulfate reductase – *aprA* AOV p-value = 0.005 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.014 *			**2**	0.028 *		
3	0.006 **	0.904		**3**	0.014 *	0.939	
4	0.417	0.288	0.126	**4**	0.884	0.097 -	0.048 *
C. Methane monooxygenase – *mmoX* AOV p-value = 0.002 **				**D. Methanol dehydrogenase – *mxaF* AOV p-value = 0.014 ***			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.008 **			**2**	0.046 *		
3	0.003 **	0.842		**3**	0.010 *	0.739	
4	0.357	0.244	0.081 -	**4**	0.327	0.749	0.501
E. NiFe hydrogenase – *hydB* AOV p-value = 0.667				**F. Methylene-tetrahydromethanopterin dehydrogenase – *mtdB* AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.975			**2**	< 0.001 ***		
3	0.979	1		**3**	< 0.001 ***	0.160	
4	0.933	0.666	0.709	**4**	< 0.001 ***	0.999	0.296
G. Heterodisulfide reductase – *hdrA* AOV p-value = 0.005 **				**H. Anaerobic carbon-monoxide dehydrogenase – *cooS* AOV p-value = 0.002 ****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.016 *			**2**	0.004 **		
3	0.004 **	0.725		**3**	0.006 **	1	
4	0.172	0.720	0.251	**4**	0.369	0.131	0.147

Appendix E.15. Shallow Basin: Summary of *p* values from AOV and *post hoc* Tukey HSD of a range of geochemical values across and between sample resolution groups. (A). Porewater manganese (Mn) concentrations; (B). Porewater iron (Fe) concentrations; (C). Porewater uranium (U) concentrations; (D). Bulk Mn values normalised by aluminium (Al); (E). Bulk Fe values normalised by Al; (F). Porewater nitrate concentrations; (G); Porewater sulfate concentrations; (H). Porewater thiosulfate concentrations; (I). Porewater phosphate concentrations; (J). Ferrozine Fe^{2+} as a fraction of total Fe content.

A. Mn (mg/L) AOV p-value = < 0.001 ***				**B. Fe (mg/L) AOV p-value = 0.510**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.855		
3	< 0.001 ***	0.112		**3**	0.590	0.905	
4	< 0.001 ***	0.023 *	0.693	**4**	0.549	0.843	0.995
C. U (μg/L) AOV p-value = 0.026 *				**D. Mn/Al (%) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.278			**2**	-		
3	0.045 *	0.474		**3**	-	-	
4	0.038 *	0.341	0.968	**4**	-	-	-
E. Fe/Al (%) AOV p-value = < 0.001 ***				**F. Nitrate (mg/L) AOV p-value = 0.977**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	-			**2**	1		
3	-	-		**3**	0.978	0.981	
4	-	-	-	**4**	1	1	0.992
G. Sulphate (mg/L) AOV p-value = < 0.001 ***				**H. Thiosulphate (mg/L) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	0.922		
3	< 0.001 ***	0.512		**3**	< 0.001 ***	< 0.001 ***	
4	< 0.001 ***	0.019 *	0.238	**4**	< 0.001 ***	< 0.001 ***	0.071 -
I. Phosphate (mg/L) AOV p-value = 0.001 **				**J. Fe^{2+} (%) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.169			**2**	< 0.001 ***		
3	< 0.001 ***	0.004 **		**3**	< 0.001 ***	1	
4	< 0.001 ***	0.002 **	0.796	**4**	< 0.001 ***	0.984	0.987

Appendix E.16. Deep Basin: Summary of *p* values from AOV and *post hoc* Tukey HSD of a range of geochemical values across and between sample resolution groups. (A). Porewater manganese (Mn) concentrations; (B). Porewater iron (Fe) concentrations; (C). Porewater uranium (U) concentrations; (D). Bulk Mn values normalised by aluminium (Al); (E). Bulk Fe values normalised by Al; (F). Porewater nitrate concentrations; (G); Porewater sulfate concentrations; (H). Porewater thiosulfate concentrations; (I). Porewater phosphate concentrations; (J). Ferrozine Fe^{2+} as a fraction of total Fe content.

A. Mn (mg/L) AOV p-value = < 0.001 ***				**B. Fe (mg/L) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.006 **			**2**	< 0.001 ***		
3	< 0.001 ***	0.141		**3**	0.048 *	0.265	
4	< 0.001 ***	0.015 *	0.653	**4**	0.920	< 0.001 ***	0.007 **
C. U (μg/L) AOV p-value = 0.002 **				**D. Mn/Al (%) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	0.009 **			**2**	< 0.001 ***		
3	0.004 **	0.926		**3**	< 0.001 ***	0.953	
4	0.657	0.089 -	0.041 *	**4**	< 0.001 ***	0.038 *	0.102
E. Fe/Al (%) AOV p-value = < 0.001 ***				**F. Nitrate (mg/L) AOV p-value = 0.309**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	< 0.001 ***			**2**	1		
3	< 0.001 ***	0.042 *		**3**	0.568	0.498	
4	< 0.001 ***	< 0.001 ***	0.068 -	**4**	0.557	0.496	1
G. Sulphate (mg/L) AOV p-value = 0.953				**H. Thiosulphate (mg/L) AOV p-value = 0.216**			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	1			**2**	1		
3	1	1		**3**	1	1	
4	0.975	0.949	0.968	**4**	0.337	0.225	0.264
I. Phosphate (mg/L) AOV p-value = < 0.001 ***				**J. Fe^{2+} (%) AOV p-value = < 0.001 *****			
Group	**1**	**2**	**3**	**Group**	**1**	**2**	**3**
2	1			**2**	< 0.001 ***		
3	0.689	0.560		**3**	< 0.001 ***	0.967	
4	0.002 **	< 0.001 ***	0.010 *	**4**	< 0.001 ***	0.947	1